ELECTRONIC DEVICES AND CIRCUITS

An Introduction

ALLEN MOTTERSHEAD
Cypress College

Prentice-Hall of India Private Limited
New Delhi-110001

ELECTRONIC DEVICES AND CIRCUITS

First Published : 1960

Printed by Jay Print Pack Private Limited, New Delhi-110015 and Published by Prentice-Hall of India Private Limited, M-97, Connaught Circus, New Delhi-110001.

CONTENTS

Preface xi

Tube and Semiconductor Notations xiii

PART 1: DIODES

Chapter 1 — Characteristics of Diodes 2

1-1 The High Vacuum Diode — Construction 2
1-2 The Semiconductor Diode — Construction 3
1-3 The Selenium Rectifier — Construction 3
1-4 The Ideal Diode 4
1-5 Diode Symbols 4
1-6 Characteristics of a High-Vacuum Diode 5
1-7 Characteristics of a Semiconductor Diode 7
1-8 Differences Between Tube and Semiconductor Characteristics 7
1-9 Diode Resistance 7
1-10 Comparison of Practical Diodes with the Ideal Diode 9
1-11 Equivalent Circuits 9
1-12 Analysis of a Circuit Containing a Diode 10
1-13 Tube Type Numbers 11
1-14 Semiconductor Numbers 12
1-15 Summary, Questions and Problems 12

Chapter 2 — Rectifying Circuits and DC Power Supplies 16

2-1 Load Line Analysis of a Diode Circuit 16
2-2 The Half-Wave Rectifier 19
2-3 Voltage Regulation 20
2-4 Ripple Factor 21
2-5 Ratio of Rectification 22
2-6 Transformer Utilization Factor 23
2-7 Disadvantages of the Half-Wave Rectifier 24
2-8 The Full-Wave Rectifier 24
2-9 The Bridge Rectifier 26
2-10 Three-Phase Full-Wave Rectifiers 26
2-11 Comparison of Rectifier Circuits 27
2-12 Reading of an AC Voltmeter 27
2-13 Summary, Questions and Problems 29

Chapter 3 — Filter Circuits for Power Supplies 33

3-1 The Inductor Filter 33
3-2 Experimental Determination of Ripple Factor 35
3-3 The Capacitor Filter 35
3-4 Ripple Factor, Approximate Method 35
3-5 Ripple Factor, More Accurate Method 38
3-6 Diode Peak Current Requirements 39
3-7 Effect of Capacitor Series Resistance 41
3-8 PIV of a Half-Wave Rectifier 41
3-9 The Choke Input or LC Filter 41
3-10 Ripple Factor in an LC Filter 42
3-11 Value of the Critical Inductance 43
3-12 Multiple LC Filters 44
3-13 The CLC or π Filter 45
3-14 Comparison of Filter Circuits 45
3-15 Summary, Questions and Problems 45

Chapter 4 — The Diode Clamper and Voltage Doubler 50

4-1 The Diode Clamper 50
4-2 Shunt Rectifier With Filter 52
4-3 Shunt Rectifier and Filter Application in a VTVM 52
4-4 Half-Wave or Cascade Voltage Doubler 54
4-5 Half-Wave Voltage Doubler Application in a VTVM 55
4-6 Full-Wave Voltage Doubler 57
4-7 Half-Wave Voltage Multiplier 58
4-8 Summary, Questions and Problems 59

Chapter 5 — Semiconductor Physics of Diodes 63

5-1 Review of Atomic Structure 63
5-2 Effect of Heat on Atomic Structure 64
5-3 Pure Germanium 64
5-4 Germanium at Room Temperature 65
5-5 Conduction in Pure Germanium 66

5-6 N-type Material 66
5-7 P-type Material 67
5-8 Conduction Processes in N-type and P-type Semiconductors 68
5-9 PN Junctions 68
5-10 Forward-Biased Junction 70
5-11 Reverse-Biased Junction 71
5-12 Effect of Temperature on PN Junction Diodes 71
5-13 Maximum Temperature Operation 73
5-14 Reverse Breakdown Voltage 74
5-15 Capacitive Effects in a PN Junction Diode 74
5-16 Dynamic Testing of a Diode 76
5-17 Summary, Questions and Problems 77

Chapter 6 — Zener Diodes, Tunnel Diodes and Thermistors 81

6-1 Zener Diodes 81
6-2 Zener Diode Specifications 81
6-3 The Voltage Regulator Circuit 82
6-4 Design of a Voltage Regulator Circuit 83
6-5 Effect of Supply Voltage Variations 84
6-6 Zener Diode Breakdown Mechanism 85
6-7 Effect of Temperature on Zener Diodes 85
6-8 Examples of Temperature Coefficient Calculation 86
6-9 Reference Zener Diodes 86
6-10 Gas-Filled Tubes 87
6-11 The Tunnel Diode 89
6-12 Applications of the Tunnel Diode 90
6-13 Thermistors 91
6-14 Externally Heated Thermistor Operation 93
6-15 Self-Heated Thermistor Operation 94
6-16 Applications in the Self-Heated Mode 94
6-17 AC Voltage Regulator 95
6-18 Summary, Questions and Problems 96

PART 2 :
GENERAL AMPLIFIER CHARACTERISTICS

Chapter 7 — General Amplifier Characteristics I: Distortion 102

7-1 Concept of Amplification 102
7-2 Amplifier Notations 103
7-3 Current Gain, A_i 103
7-4 Voltage Gain, A_v 103
7-5 Power Gain, A_p 103
7-6 Amplifier Input Resistance, R_i 104
7-7 Amplifier Output Resistance, R_o' 105
7-8 Maximum Power Transfer 106
7-9 Conversion Efficiency 106
7-10 Classes of Amplifier Operation 107
7-11 Harmonic Distortion 108
7-12 Three-Point Method of Calculating Harmonic Distortion 108
7-13 Five-Point Method of Calculating Harmonic Distortion 110
7-14 Oscilloscope Display of an Amplifier's Dynamic Transfer Curve 111
7-15 Measurement of Harmonic Distortion 114
7-16 Other Types of Amplifier Distortion 115
7-17 Summary, Questions and Problems 116

Chapter 8 — General Amplifier Characteristics II: Decibels 120

8-1 Decibels 120
8-2 Other Equations for Decibel Computation 121
8-3 Zero Decibel Reference Level 122
8-4 Use of a Voltmeter as a Decibel Indicator 123
8-5 Voltmeter Range Correction Factors 123
8-6 Impedance Correction Factor 124
8-7 Frequency Response Curves 124
8-8 Amplifier Bandwidth 125
8-9 Use of an Output Power Meter 126
8-10 Phase Relationships in Amplifiers 128
8-11 Square-Wave Testing 129
8-12 Hearing Response, Loudness Levels 129
8-13 Soundlevel Meters 132
8-14 Amplifier Tone Controls 132
8-15 RIAA Equalization 134
8-16 Miscellaneous Amplifier Specifications 135
8-17 Loudspeaker Requirements 136
8-18 Summary, Questions and Problems

PART 3 :
BIPOLAR TRANSISTOR AMPLIFIERS

Chapter 9 — Basic Characteristics of the Transistor 142

9-1 Basic Transistor Amplifier 142
9-2 Two-Diode Analogy for a Transistor 143
9-3 Transistor Input Characteristics 144
9-4 Transistor Collector Characteristics, CE 144
9-5 Collector Cutoff Current, I_{CEO} 145
9-6 Forward Current Transfer Ratio, CE 145
9-7 Permissible Operating Area of a Transistor, CE 147
9-8 The Basic Common-Base Amplifier, CB 147
9-9 Forward Current Transfer Ratio, CB 148
9-10 Relation Between α and β 149
9-11 Collector Cutoff Current, I_{CBO} 150
9-12 Transistor Testers 151
9-13 Physical Explanation of CB and CE Amplifying Action 151
9-14 Reduction of CE Leakage Current to I_{CO} 154
9-15 Common-Collector Amplifier 155
9-16 PNP and NPN Transistor Amplifiers 155
9-17 Transistor Construction 155
9-18 Identifying the Transistor Leads 157
9-19 Summary, Questions and Problems 157

Chapter 10 — The Common-Base Amplifier 161

10-1 Analysis of the Common-Base Amplifier 161
10-2 The Amplifier as a Whole 165
10-3 Power Relationships 168
10-4 Efficiency 168
10-5 Phase Relationships in a CB Amplifier 169
10-6 Input Waveform Considerations 170
10-7 Harmonic Distortion 170
10-8 Summary, Questions and Problems 170

Chapter 11 — The Common-Emitter Amplifier 175

11-1 Graphical Analysis of a Common-Emitter Class A Amplifier 175
11-2 Input and Output Resistance 177
11-3 Effect of Adding an AC Load 178
11-4 Conversion Efficiency of a Class A Amplifier with a Direct-Coupled Resistive Load 178
11-5 Phase Relationships in a CE Amplifier 180
11-6 Input Waveform Considerations 180
11-7 Harmonic Distortion Calculation 182
11-8 Oscilloscope Display of Characteristic Curves 185
11-9 Comparison of Basic Transistor Amplifiers 186
11-10 Summary, Questions and Problems 186

Chapter 12 — Thermal Stability I: Transisto Biasing 192

12-1 Factors Contributing to Thermal Instability 192
12-2 Effect of Temperature Increase 193
12-3 Stability Factor S 193
12-4 Common-Base Stability 194
12-5 Collector-to-Base Bias 195
12-6 Disadvantages of Collector-to-Base Bias 196
12-7 Emitter Bias 196
12-8 Voltage Divider Bias with Emitter Bias 199
12-9 Emitter Bypass Capacitor 202
12-10 Summary of Stabilization Circuits 202
12-11 Additional Stability Factors 202
12-12 Bias Compensation 203
12-13 Summary, Questions and Problems 204

Chapter 13 — Thermal Stability II: Transistor Dissipation 209

13-1 Transistor Dissipation 209
13-2 Significance of the Quiescent Point in Thermal Runaway 210
13-3 Thermal Resistance 211
13-4 Calculation of Transistor Dissipation 213
13-5 Heat Sink Theory 213
13-6 Heat Sink Hardware and Thermal Characteristics 215
13-7 Heat Sink Mounting 217
13-8 Selection of Heat Sink Size 218
13-9 Conditions for Thermal Stability 220
13-10 Thermal Fatigue 222
13-11 Summary, Questions and Problems 223

Chapter 14 — Hybrid Equivalent Circuit for a Transistor 229

14-1 Conversion of an Amplifier Circuit to a Standard Form 229
14-2 General "Black Box" Theory 230
14-3 Hybrid h-Parameters 231
14-4 Obtaining the Hybrid h-Parameters 231

14-5 Typical h-Parameter Values 233
14-6 Amplifier Equations 237
14-7 Voltage and Current Gains Taking Into Account R_g of Source 238
14-8 Dual Matching in an Amplifier for Maximum Power Gain 240
14-9 Dependence of Amplifier Characteristics on R_L and R_g 241
14-10 Comparison of CB, CE and CC Characteristics 241
14-11 Summary, Questions and Problems 244

Chapter 15 — Low-Frequency Response of the Transistor Amplifier 248

15-1 Effect of an Emitter Bypass Capacitor on Low-Frequency Response 248
15-2 Effect of a Coupling Capacitor on Low-Frequency Response 252
15-3 Cascading of CE Stages, Mid-Frequency Gains 253
15-4 Low-Frequency Response of Cascaded Stages 255
15-5 Amplifier Low-Frequency Response to a Square Wave 256
15-6 Transformer-Coupled Transistor Amplifiers 258
15-7 Low-Frequency Response of a Transformer-Coupled Amplifier 260
15-8 Step Response of a Transformer-Coupled Amplifier 261
15-9 Summary, Questions and Problems 262

Chapter 16 — High-Frequency Response of the Transistor Amplifier 266

16-1 High-Frequency Model for the Common-Emitter Amplifier 266
16-2 Approximate CE High-Frequency Model with a Resistive Load 267
16-3 CE Short-Circuit Current Gain 268
16-4 High-Frequency Current Gain with a Resistive Load 270
16-5 High-Frequency Response of Cascaded CE Stages 273
16-6 Amplifier High-Frequency Response to a Square Wave 274
16-7 High-Frequency Response of a Transformer-Coupled Amplifier 275
16-8 Transistor Noise 277
16-9 Summary, Questions and Problems 279

Chapter 17 — Negative Feedback in Transistor Amplifiers 282

17-1 General Theory of Feedback 282
17-2 Reasons for Negative Feedback 284
17-3 Loop Gain 288
17-4 Types of Negative Feedback in Transistor Circuits 289
17-5 The Darlington Connection 294
17-6 Biasing the Darlington Amplifier 296
17-7 Summary, Questions and Problems 298

PART 4 : OSCILLATORS

Chapter 18 — Transistor Oscillators and Multivibrators 304

18-1 Effect of Positive Feedback 304
18-2 Requirements for Oscillation 305
18-3 The Phase-Shift Oscillator 306
18-4 Another Version of the Phase-Shift Oscillator 307
18-5 Wien Bridge Oscillator 307
18-6 Resonant Circuit Oscillators 307
18-7 The Maximum Frequency of Oscillation of a Transistor 310
18-8 Decoupling Filters 310
18-9 The Transistor as a Switch 312
18-10 Switching Times in a Transistor 314
18-11 Multivibrators 315
18-12 Summary, Questions and Problems 320

PART 5 : TRANSISTOR POWER AMPLIFIERS

Chapter 19 — Transistor Power Amplifiers I: Class A 326

19-1 Class A Direct-Coupled Resistive Load 326
19-2 Transformer-Coupled Resistive Load 327
19-3 Design Theory 328
19-4 Power Amplifier Design 329
19-5 Harmonic Distortion 331
19-6 Power Output 332
19-7 Variation of Output Power with Load 332
19-8 Thermal Runaway 332
19-9 Output Transformer Saturation 333

19-10 Disadvantages of a Single-Ended Transformer-Coupled Amplifier 333
19-11 Push-Pull Amplifiers 333
19-12 Description of Operation of a Class A Push-Pull Amplifier 334
19-13 Theory of Operation of a Class A Push-Pull Amplifier 335
19-14 Summary, Questions and Problems 336

Chapter 20 — Transistor Power Amplifiers II: Class B 339

20-1 The Class B Push-Pull Amplifier 339
20-2 Crossover Distortion 340
20-3 Class AB Push-Pull Amplifier 340
20-4 Transistor Phase Inverter 341
20-5 Conversion Efficiency of a Class B Amplifier 341
20-6 Relation Between Maximum Output Power and Load Resistance 343
20-7 Relation Between Maximum Output Power and Transistor Dissipation 344
20-8 Design of a Class B Push-Pull Amplifier 345
20-9 Other Class B Push-Pull Amplifiers 347
20-10 Complementary Symmetry 347
20-11 Practical Complementary Symmetry Amplifier 348
20-12 Summary, Questions and Problems 350

PART 6 : FIELD EFFECT TRANSISTORS

Chapter 21 — Field-Effect Transistor Amplifiers 356

21-1 Advantages and Disadvantages of the FET 356
21-2 Basic Construction of the JFET 357
21-3 Characteristic Curves of the JFET 358
21-4 Principles of Operation of the JFET 359
21-5 Effect of V_{DS} on Channel Conductivity 359
21-6 Channel Ohmic Region and Pinch-Off Region 360
21-7 Characteristic Parameters of the FET 361
21-8 Effect of Temperature on FET Parameters 362
21-9 Common-Source AC Amplifier 364
21-10 Fixed Bias with Self-Bias 366
21-11 The Common-Drain or Source Follower 368
21-12 The Common-Gate FET Amplifier 370
21-13 Frequency Response of the FET Amplifier 370
21-14 Other Amplifier Configurations 372
21-15 Summary, Questions and Problems 372

Chapter 22 — MOSFETs and Other Applications of FETs 378

22-1 The Depletion MOSFET 378
22-2 The Enhancement MOSFET 380
22-3 Differences Between JFETs and MOSFETs 381
22-4 Handling Precautions for MOSFETs 381
22-5 Dual-Gate MOSFETs 382
22-6 Integral Gate Protection 383
22-7 Testing Field-Effect Transistors 384
22-8 Application of a Dual-Gate MOSFET in an AGC Amplifier 385
22-9 Other Applications of FETs 387
22-10 The Field-Effect Diode 391
22-11 Summary, Questions and Problems 393

PART 7 : INTEGRATED CIRCUITS

Chapter 23 — Linear Integrated Circuits: Operational Amplifiers 400

23-1 The Operational Amplifier 400
23-2 The Inverting Differential Operational Amplifier 401
23-3 Analog Computer Solution of Linear Equations 403
23-4 Increasing Input Impedance in an Inverting Amplifier 405
23-5 The Noninverting Differential Operational Amplifier 406
23-6 The Differential Amplifier 407
23-7 A General Purpose IC Operational Amplifier 408
23-8 Common-Mode Rejection Ratio 409
23-9 Emitter-Coupled Differential Amplifier 410
23-10 High-Performance Operational Amplifier 412
23-11 Increased Differential Input Impedance 413
23-12 Applications of Differential Operational Amplifiers 413
23-13 IC Audio Power Amplifier 419
23-14 Summary, Questions and Problems 419

Chapter 24 — An Introduction to the Fabrication of Integrated Circuits 426

24-1 Evolution of Integrated Circuits 426
24-2 The Monolithic Integrated Circuit 427
24-3 Integrated Circuit Components 432
24-4 Methods of Fabricating Integrated Circuits 435
24-5 Complementary Symmetry MOS Integrated Circuit COS/MOS 438
24-6 Large-Scale Integration (LSI) 440
24-7 Summary, Questions and Problems 440

Chapter 25 — NonLinear Integrated Circuits I: Combinational Digital Circuits 446

25-1 The Binary Concept 446
25-2 Basic Logic Gates 448
25-3 De Morgan's Law 451
25-4 NAND and NOR Logic Gates 452
25-5 Comparison of Logic Families 461
25-6 Implementing Logic Circuits with NAND-NOR Gates 462
25-7 The Meaning of Logic Implementation 463
25-8 Summary, Questions and Problems 464

Chapter 26 — NonLinear Integrated Circuits II: Sequential Digital Circuits 471

26-1 The Basic R-S Flip-Flop 471
26-2 Clock Pulses 473
26-3 The Clocked R-S Flip-Flop 473
26-4 The Clocked J-K Flip-Flop — DTL, TTL 474
26-5 The Clocked J-K Flip-Flop — RTL 475
26-6 Other Types of Flip-Flops 477
26-7 Clock Pulse Requirements 478
26-8 The Monostable Multivibrator (One-Shot) 478
26-9 The Schmitt Trigger 480
26-10 Applications of Flip-Flops in Binary Counters 483
26-11 Some Basic Applications of Counters 486
26-12 Summary, Questions and Problems 490

PART 8: INDUSTRIAL DEVICES

Chapter 27 — Optoelectronic Devices 498

27-1 Fundamentals of Light 498
27-2 Photoconductive Sensors 503
27-3 Applications of Photodiodes and Phototransistors 509
27-4 Photovoltaic Sensors 513
27-5 Photoemissive Sensors 515
27-6 Light Emitters 515
27-7 AlphaNumeric Displays 516
27-8 Photocouplers 517
27-9 Summary, Questions and Problems 518

Chapter 28 — Thyristors and the Unijunction Transistor 524

28-1 The Silicon Controlled Rectifier (SCR) 524
28-2 The Triac 532
28-3 Triggering Devices 533
28-4 Typical Diac-Triac Phase-Control Circuit 534
28-5 The Unijunction Transistor (UJT) 536
28-6 Summary, Questions and Problems 541

Appendix A — Basic Characteristics of the Triode Amplifier 546

A-1 Triode Construction 546
A-2 Triode Characteristics 547
A-3 Tube Properties 548
A-4 Relationship Between μ, r_p, g_m 551
A-5 Permissible Area of Operation 551
A-6 Operation of the Triode as a DC Amplifier 552
A-7 Theory of Operation of an AC Amplifier: Bias 554
A-8 Gain of an AC Amplifier 555
A-9 Maximum Output Voltage Available from an Amplifier 558
A-10 Use of a Blocking Capacitor 558
A-11 Equivalent Circuit for a Small Signal AC Amplifier — Grounded Cathode 559
A-12 Summary, Questions and Problems 560

Appendix B — Analysis of Class A Triode Amplifiers 564

B-1 Evolution of Self Bias 564
B-2 Bypass Capacitor 565
B-3 Grid-Leak Resistor 566

B-4 Determination of Operating Point 566
B-5 Bias Line 566
B-6 AC Load Line 567
B-7 Effect of an AC Load 569
B-8 Effect of Coupling Capacitor 570
B-9 Performance Calculations for a Class A Triode Amplifier 571
B-10 Other Amplifier Configurations 575
B-11 Cathode Follower 575
B-12 Grounded-Grid Amplifier 576
B-13 Summary, Questions and Problems 577

Appendix C — Frequency Response of Cascaded Triode Amplifiers 582

C-1 Direct Coupled Amplifier 582
C-2 The Triode at High Frequencies 583
C-3 Triode Input Capacitance 584
C-4 Miller Effect 585
C-5 Resistance-Capacitance (RC) Coupled Amplifier 586
C-6 Frequency Response Curve 588
C-7 Gain-Bandwidth Product 588
C-8 Negative Input Resistance 591
C-9 Frequency Response of a Cathode Follower 592
C-10 Summary, Questions and Problems 593

Appendix D — Pentode Voltage and Power Amplifiers 598

D-1 The Tetrode 598
D-2 The Pentode 599
D-3 Voltage Gain of a Pentode Amplifier 602
D-4 Graphical Analysis of a Pentode Voltage Amplifier 602
D-5 High Frequency Response of a Cascaded Pentode Voltage Amplifier 603
D-6 Low Frequency Response of a Cascaded Pentode Voltage Amplifier 604
D-7 Radio-Frequency (RF) Voltage Amplifiers 604
D-8 Power Amplifiers 606
D-9 The Triode Cascode Amplifier 608
D-10 Limitations of Amplifiers Using Tubes 609
D-11 Summary, Questions and Problems 609

Appendix E — Phase-Shift Measurement 613

Appendix F — Abbreviated Math Tables 615

Answers to Odd-Numbered Problems 625

Index 635

PREFACE

Electronic Devices and Circuits — An Introduction is written for students in electronics technology at a junior college, state college, or technical institute. It is designed to follow a semester of ac and dc circuit theory and presumes no previous knowledge of tubes or transistors. The only mathematics necessary is intermediate algebra and a little trigonometry. Although the rate of change concept is used occasionally, as in the determination of the thermal stability of a transistor amplifier, calculus is not necessary at all.

The author believes that a text covering theory of practical devices and circuits should be accompanied by a closely correlated laboratory manual. *Laboratory Experiments for Electronic Devices and Circuits* was written for this purpose. Together, these two books cover enough material for a two semester sequence, comprising at least 3 hours of lecture and from 3 to 6 hours laboratory per week.

Regarding the material, much has been written about the "obsolescence" of tubes. Some instructors feel that all tube theory should be removed from the curriculum, and everything become solid-state oriented. Admittedly, industry has become almost 100% solid-state in its manufacture of equipment, hardware, and appliances, but no one can deny that a tremendous amount of tube equipment is still in service, not to mention the fact that the FCC license examinations still require some tube theory.

Accordingly, four chapters are available in the appendix and are devoted to tubes and their characteristics. For those who use this text as a two-semester sequence, these chapters may be used as a vehicle to put across many concepts which are used in transistor work. At least initially, the tube seems to be easier to describe and understand from a physical standpoint, and it can be used to introduce the analytical methods of graphical analysis used with the transistor.

However, for those using the text as a one or two-semester solid-state course, discussion of the transistor amplifier without benefit of first considering tube amplifiers is made easier by two chapters on general amplifier characteristics. Chapters recommended for a one-semester course are: 1, 2, 3, 5, 7, 8, 9, 11, 12, 14, 18, Sections 19-12 and 13, 20, 21, 23, 24, 25, 27. Chapter 24 could be assigned for outside reading if time is limited.

The notation used for electron and semiconductor devices conforms to the standard symbols recommended by the IEEE. An explanation of the notation is included and should be constantly referred to by the student as he reads the text. Accordingly, conventional current flow has been used throughout, although electron flow is used in the tube discussions, and of course both electron and hole flow in the interior of the transistor. This may be a questionable decision, but it seems that the solid-state symbols logically lend themselves to the conventional current flow direction while the electron flow should be reserved for the electron tube. However, the qualified technician should be well versed in both, since they really are dialects of the same language.

A recurring pattern throughout the text is the presentation of a device, a discussion of its characteristics, and then consideration of its equivalent circuit or mathematical model. This is then followed by typical circuit applications. The h-parameters for a transistor's hybrid equivalent circuit are introduced primarily for their use in the discussions of frequency response and negative feedback in transistor amplifiers.

This text does not cover applications in the fields of communications or computers (except for some introductory logic circuits), since it is usually necessary to have separate detailed courses on these subjects. It is hoped, however, that this book will provide a firm foundation for these topics by covering an important link between basic electricity and advanced electronics applications.

I wish to express my appreciation to the following companies who made available photographs of instruments and devices used in this text: Hewlett-Packard, General Radio, Tektronix, Hickok Electrical Instrument, and RCA.

Finally, I would like to thank my mother who painstakingly typed most of the manuscript, and my family whose sacrifices made the whole project possible.

TUBE AND SEMICONDUCTOR NOTATIONS

The following notation corresponds to that adopted by the IEEE (Institute of Electrical and Electronic Engineers) in 1963-1964. It is to be noted that no longer is e or E used for a source of voltage or emf, but instead v or V. This is to help reduce confusion when dealing with the emitter (e or E) in a transistor. Those who are familiar with the old notation will no longer find Ebb or B+ used for tube supply voltages but instead V_{PP}, to conform to the transistor notation. However, once the notation is learned it provides a precise method for keeping track of ac and dc quantities.

1. The electrodes of a tube or transistor are represented by uppercase letters (G for grid, P for plate, K for cathode, B for base, C for collector, E for emitter).

2. If an electrical quantity varies with time its *instantaneous* value is represented by a lowercase letter (i for current, v for voltage, p for power).

3. Uppercase letters are used for the average, maximum and effective (rms) values of time varying quantities as well as for pure dc (I, V and P).

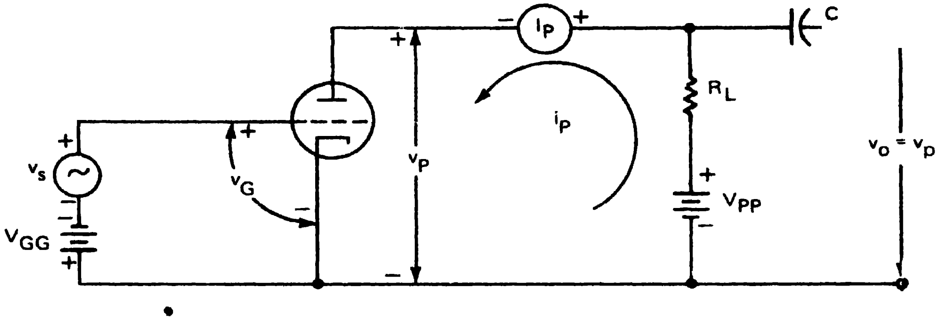
4. Subscripts are used on the above letters according to the electrode involved and the reference level. If an average (dc) or a total quantity (consisting of an average level plus a time varying component) is referred to (reference level is zero), uppercase letters are used (G,P,K,B,C,E). If only the time varying component is of interest (quiescent value is refer ence) lowercase letters are used (g,p,k,b,c,e).

5. The magnitude of a supply voltage is indicated by a double subscript according to the electrode involved (GG,PP,KK,BB,CC,EE).

6. Additional subscripts of *m* and *avg* may be used to distinguish between maximum and average values. Otherwise an rms value is understood.

7. The double subscript notation is used for voltages in a transistor only if the reference electrode is not clearly understood. To help explain this notation, the typical circuits and graphs on the following pages are offered. You should study them carefully and refer to them constantly while reading the text until you are completely familiar with the notation.

(a) Typical tube amplifier:

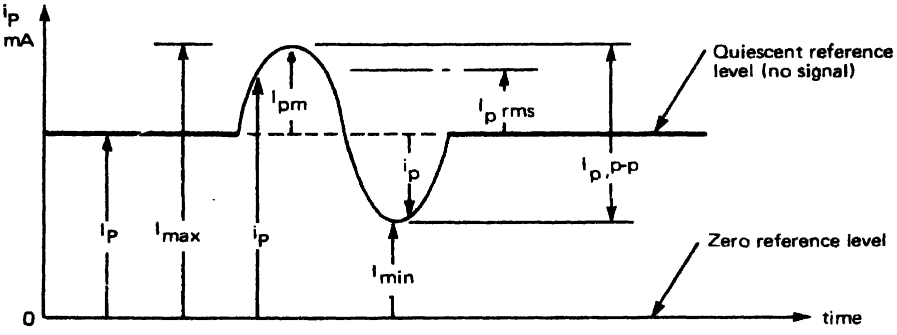


$$v_s = V_{sm}\sin \omega t = \sqrt{2}\, V_s \sin \omega t$$

$$v_G = v_s - V_{GG} = V_{sm}\sin \omega t - V_{GG}$$

$$v_P = V_{PP} - i_P R_L .$$

Consider a graph of total plate current i_P:

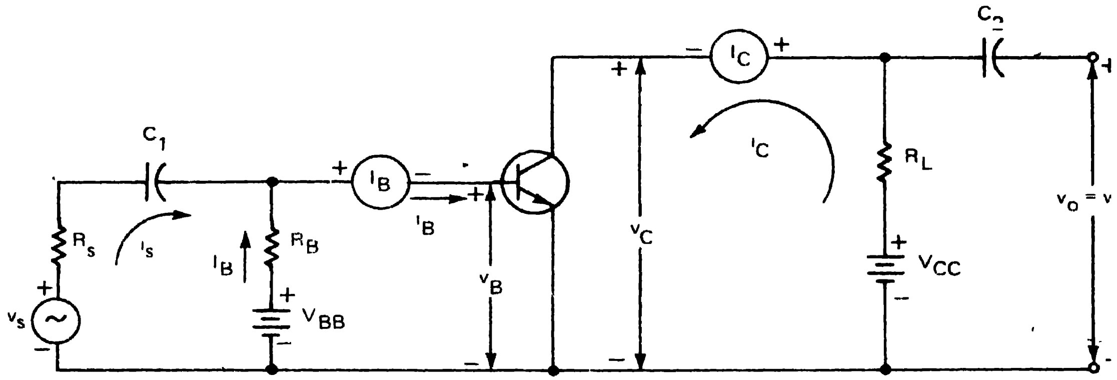


$$i_p = I_{pm}\sin \omega t = \sqrt{2}\, I_p \sin \omega t$$

$$i_P = I_P + i_p$$

$$I_p \text{ peak-to-peak} = I_{p,p\text{-}p} = I_{max} - I_{min}$$

(b) Typical transistor amplifier:



$$i_s = I_{sm}\sin\omega t = \sqrt{2}\, I_s \sin\omega t$$

$$i_B = I_B + i_s = I_B + I_{sm}\sin\omega t$$

$$v_{CE} = V_{CC} - i_C R_L = v_{ce} + V_{CE} \quad \text{or}$$

$$v_C = V_{CC} - i_C R_L = v_c + V_C$$

$$i_C = I_C + i_c$$

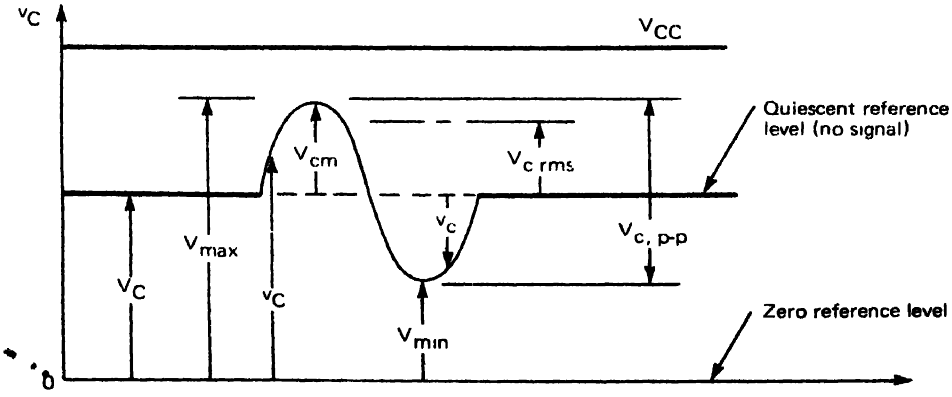
$$i_c = I_{cm}\sin\omega t = \sqrt{2}\, I_c \sin\omega t$$

$$V_{ce} \text{ peak-to-peak} = V_{c,p\text{-}p}$$

$$V_{max} - V_{min}$$

The double subscript notation is not necessary since it is clear that the emitter is the reference electrode.

Consider a graph of the total collector voltage (with respect to emitter):



DIODES

PART 1

1

CHARACTERISTICS OF DIODES

1-0. INTRODUCTION

Diodes are to be found in all phases of electronics, radio and TV, industrial electronics, computers, etc. Diodes are used extensively in power supplies to convert ac into dc. In this application they are referred to as rectifiers. However, their many properties, especially those of the semiconductor devices, make it very important to properly understand the various types of diodes and their characteristics. Although the primary feature of the diode is to allow the flow of current in one direction only, diodes can be used in a variety of nonrectifying circuits. Some photodiodes can even emit light when current passes through them, promising a solid-state TV or oscilloscope screen in the future.

The high-vacuum diode has largely been replaced by silicon and selenium rectifiers, although the vacuum tube still remains where high inverse voltages are found, such as in X-ray equipment high-voltage power supplies. However, where cost is no objection, many high-voltage power supplies can be made fully solid-state.

In this chapter the ideal diode is introduced, and then the characteristics of practical diodes are examined to see how closely they approach the ideal device. The concept of an equivalent linear circuit for a diode is also introduced, and its aid in analyzing the operation of a device in a circuit is demonstrated

1-1. THE HIGH VACUUM DIODE – CONSTRUCTION

All tubes, whether diode, triode, tetrode or pentode, have two items in common: the anode and cathode. The symbol used for a vacuum tube diode shows these two elements as in Fig. 1-1.

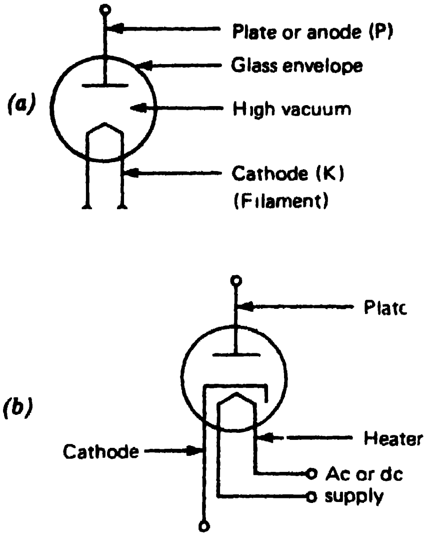


Fig. 1-1. Symbol of a vacuum diode (a) with a directly heated cathode; (b) with an indirectly heated cathode.

Methods of Causing Emission

The function of the cathode is to emit electrons, which are then attracted to the positive plate or anode. There are three methods of causing emission of electrons from the cathode:

(a) Thermionic.
(b) Photoelectric.
(c) Electric field.

Thermionic emission is the most widely used. In fact, the vacuum diode is often referred to as a thermionic diode because it relies upon heat for electron emission.

Thermionic emission may employ either a directly or an indirectly heated cathode. The former type is used in tubes primarily designed to be operated from batteries, and the cathode is often called the filament. The indirectly heated cathode type has a separate heater which is usually found in ac operated tube equipment. Its symbol is shown in Fig. 1-1(b). Although in many schematics the heater is omitted for clarity it is understood to be connected to a suitable supply.

Size of Vacuum Diodes

The size of a diode will obviously depend upon its current and voltage capability. A 5Y3-GT duodiode (containing two plates and a directly heated common cathode), capable of a maximum dc output current of 125 mA, is 1-1/4 inches in diameter and 3-1/4 inches in length, including the pins. By contrast, a 6AL5, an indirectly heated duodiode (containing two plates and two cathodes), is only 1-3/4 inches long and 3/4 inch in diameter. Its maximum dc current per plate is 9 mA.

1-2. THE SEMICONDUCTOR DIODE – CONSTRUCTION

The detailed construction and operation of a semiconductor diode will be covered in a later chapter, but for purposes of size comparison and symbol recognition the following should suffice.

A typical rectifying diode has silicon (or germanium for special applications) as its basic material. One end is doped (modified) to provide an increase in electron availability compared with the pure silicon and is then known as an N-type material. The other end is doped to reduce the electron availability and is referred to as a P-type material. A P-N junction is thus formed in an overall size of approximately a 1/8-inch cube. See Fig. 1-2(a). This is then encapsulated in some sealing compound after connecting wires are attached. The body of the diode, for a 1 A capability, will seldom exceed one quarter of an inch in diameter and half an inch in length.

The main point to note is the lack of a separation in the device between the two terminals, such as exists in a thermionic diode, and its obvious compactness. How such a continuous mass can allow current flow in one direction but prevent it in the other will be an interesting discussion to be covered in Chapter 5.

The symbol used for the semiconductor diode (and sometimes used for a thermionic diode) is shown in Fig. 1-2 along with the thermionic diode for comparison.

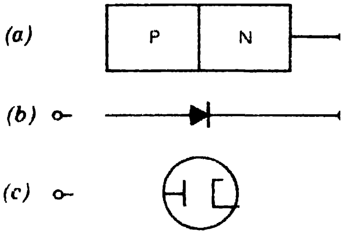


Fig. 1-2. (a) Sketch of P-N junction.
(b) Symbol for semiconductor diode.
(c) Symbol for thermionic diodes.

1-3. THE SELENIUM RECTIFIER – CONSTRUCTION

The selenium rectifier, also known as a dry-disk metallic rectifier, operates on a similar principle to the P-N junction semiconductor diode, but is much bulkier. The basic selenium cell consists of a nickel plated aluminum base plate (anode) coated with selenium, over which a low temperature alloy is sprayed to form the cathode. Heat treatment of the cell forms a "barrier layer" in the selenium which is farther widened during manufacture by the application of a thin layer of insulation to increase reverse voltage capability. A sketch of a single cell is shown in Fig. 1-3 along with the diode symbol to indicate polarity.

Since the reverse voltage capability of a single cell is only 28 V dc, many cells are stacked to provide the required working voltage. The cross-sectional area of the cells determines how much current can be handled, and cooling fins are attached to increase heat dissipation. Cells may vary in size from one quarter of an inch square to six inches square. A diode capable of handling 75 mA dc with a peak inverse voltage of 500 V would typically occupy a one inch cube.

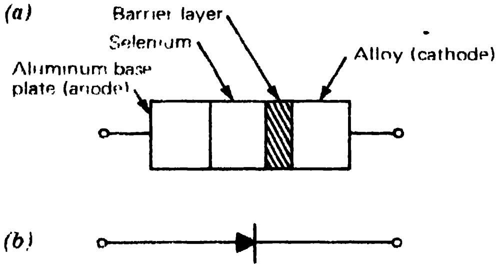


Fig. 1-3. Sketch showing: (a) the construction of a selenium diode; (b) the corresponding diode symbol.

Selenium rectifiers are bulkier than silicon diodes, but they are rugged and have the advantage of being able to withstand short time overloads or surges which are very harmful to silicon diodes. Accordingly, they are very popular in elevator control panels, but they are just as at home in a cyclotron application which may require 50,000 volts at 5 milliamperes or in electroplating equipment which may require 100,000 amperes at 50 volts dc.

1-4. THE IDEAL DIODE

We have examined briefly the construction of three types of diodes and have found that the semiconductor diode is the most compact of all three. However, little consideration has been given to their electrical characteristics. To make a comparison as to which diode is the best electrically, let us consider what characteristics the ultimate or ideal diode would have if it were possible to construct one.

What we want is a two-terminal device that is polarity sensitive. That is, the diode will allow current to flow through it without any resistance with one polarity of applied voltage, but when the voltage is reversed the diode will not permit any current to flow. This may be represented by the V-I characteristics shown in Fig. 1-4. Also shown is the polarity of voltage applied to the diode and the corresponding electrical action represented by a switch.

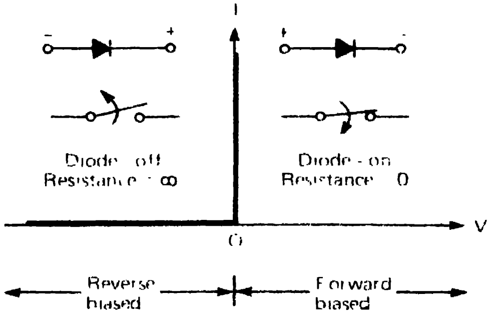


Fig. 1-4. V-I characteristic curves of an ideal diode.

When the diode is "off," or providing a blocking function, it is said to be *reverse-biased*, while a *forward-biased* condition means the diode is "on" and permits current flow. (The actual amount of current that will flow through the diode is determined by the applied voltage and the total resistance in series with the diode.) In this respect the ideal diode is a switch: either open or closed, depending upon the polarity of the voltage placed across it. We will use the symbol in Fig. 1-5(a) to represent an ideal diode.

1-5. DIODE SYMBOLS

As a summary of the symbols used for the different diodes and the direction of current flow associated with them, refer to Fig. 1-5. It should be noted that the arrow in the semiconductor symbol points in the same direction as the conventional current flow in the forward-biased condition.

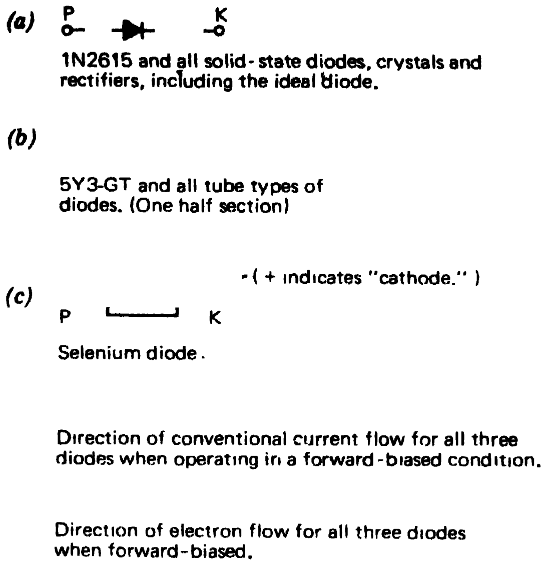

(a)
P
K
1N2615 and all solid-state diodes, crystals and rectifiers, including the ideal diode.
(b)
5Y3-GT and all tube types of diodes. (One half section)
(+ indicates "cathode.")
(c)
P
K
Selenium diode.
Direction of conventional current flow for all three diodes when operating in a forward-biased condition.
Direction of electron flow for all three diodes when forward-biased.

Fig. 1-5. Symbols for: (a) Semiconductor diode; (b) Thermionic diode; (c) Selenium diode; and current and electron flow for all three diodes.

1-6. CHARACTERISTICS OF A HIGH-VACUUM DIODE*

The characteristic curve is a graphical picture of how the current and voltage are related to each other in the device, and can be obtained for a diode from the circuit shown in Fig. 1-6.

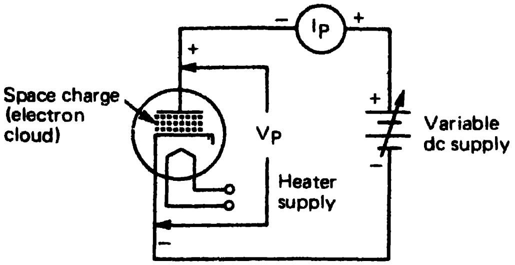

Ip
−
+
+
Space charge (electron cloud)
Vp
+
Variable dc supply
−
Heater supply
−

Fig. 1-6. Circuit used to obtain the characteristics of a diode.

As the supply voltage is increased, the plate current flowing through the tube, I_p, also increases, since the negative electrons are attracted to the positive plate, but in a nonlinear manner. That is, the voltage across the tube, V_P, is not directly proportional to the current, I_P. It is for this reason that characteristic curves are drawn. A typical vacuum diode characteristic curve of I_P vs. V_P is plotted in Fig. 1-7.

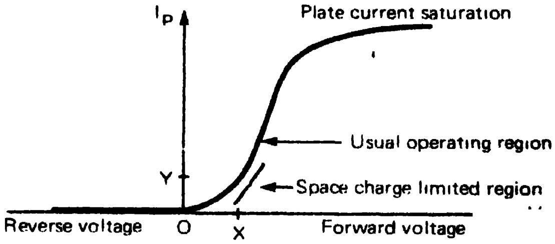


Fig. 1-7. Typical characteristic curve of a vacuum diode.

Note that the *independent* variable, the voltage that is applied to the tube, is plotted horizontally. This is always the case in any graphical presentation. The *dependent* variable, the current that flows as a *result* of the applied voltage, is plotted vertically.

In the space charge region (see Fig. 1-6), a cloud of electrons forms between cathode and plate, creating a "virtual cathode." It is from this cloud that electrons are attracted to the positive plate. Under conditions of *plate current saturation*, an increase in plate voltage causes very little increase in plate current because all the electrons that the cathode can emit are being collected by the plate.

***This section may be omitted with no loss of continuity.**

Peak Inverse Voltage

Reversing the polarity of the supply voltage and again increasing its magnitude will produce no current flow through the tube, since the plate is cold and emits no electrons. This is the case provided the peak-inverse-voltage (PIV) is not exceeded, for example, 1400 V for a 5Y3-GT. The PIV is determined by the separation of the electrodes and is the highest voltage that may be applied in the reverse direction without the air gap breaking down and providing conduction. This is called flashover and ranges up to 200,000 V for tubes used in X-ray equipment.

Typical Operating Voltages

To provide some idea of the magnitude of voltage and current involved in a tube, consider point (X, Y) in Fig. 1-7, at which the extreme curvature of the characteristic curve tends to become more linear. For a 6AL5, it is at approximately X = 4 volts and Y = 15 mA; for a 5Y3-GT, at X = 40 volts and Y = 75 mA. The tube is usually operated, for rectifying purposes in power supplies, near the lower part of the linear portion of the curve, to avoid too big a voltage drop occurring across the tube.

Low Voltage Operation

In actual practice the characteristic curve may not pass through the origin but slightly above it. That is, when the plate voltage is zero but a complete path exists between plate and cathode through a sensitive ammeter, a small current in the order of a few microamperes may be detected. This is due to the initial velocity of the electrons emitted from the cathode having enough kinetic energy to make it across the gap to the plate. The loss of electrons from the cathode makes it (and the plate to which it is connected) slightly positive thus allowing a small current. However, this current is so small that for practical purposes it can be ignored.

Dependence Upon Temperature

The characteristic curve is also dependent upon the cathode temperature, so that the curve "shifts down" and plate current saturation will occur at lower plate voltage and plate current if there is a reduction in cathode heating.

Child's Law

One of the purposes of this chapter is to acquaint the reader with mathematical models and equivalent circuits of active devices. In the space charge region of the V-I characteristic curve it is possible to relate the variables by Child's Law:

$$I_P = K\,V_P^{\frac{3}{2}} \tag{1-1}$$

where K is a constant of the tube or diode and can be found experimentally. Child's Law is valid only for tubes. K is usually determined for V_P in volts and I_P in mA. From the V-I curve a particular value of V_P (chosen conveniently) and the corresponding value of I_P would permit the determination of K. This would yield an equation from which I_P could be found mathematically for any given value of V_P, *without* reference to a curve or experimentation, in the region for which the equation is valid.

Example 1-1. The V-I characteristic curve of a vacuum diode passes through $V_P = 49$ volts and $I_P = 81$ mA in the vicinity of the space charge region. Determine: (a) An equation relating I_P to V_P in the space charge region. (b) The plate current expected for a voltage drop of 64 volts across the tube. (c) The voltage across the diode when a current of 27 mA is flowing.

Solution.

(a) $I_P = K\,V_P^{\frac{3}{2}} \quad \therefore K = \dfrac{I_P}{V_P^{\frac{3}{2}}} = \dfrac{81}{49^{\frac{3}{2}}} = \dfrac{81}{(\sqrt{49}\,)^3}$

$$= \frac{81}{7^3} = \frac{81}{343} \doteq 0.236$$

$\therefore I_P = 0.236\, V_P^{\frac{3}{2}}$, I_P in mA, V_P in volts.

$$\text{(b) } I_P = 0.236\, V_P^{\frac{3}{2}} = 0.236 \times 64^{\frac{3}{2}}$$

$$= 0.236 \times 512 = 121 \text{ mA.}$$

$$\text{(c) } 0.236\, V_P^{\frac{3}{2}} = I_P \qquad \therefore V_P^{\frac{3}{2}} = \frac{I_P}{0.236}$$

$$\therefore V_P = \left(\frac{I_P}{0.236}\right)^{\frac{2}{3}} = \left(\frac{27}{0.236}\right)^{\frac{2}{3}} = 114.4^{\frac{2}{3}}$$

$$= \left(\sqrt[3]{114.2}\right)^2 = 4.85^2 = 23.6 \text{ volts.}$$

1-7. CHARACTERISTICS OF A SEMICONDUCTOR DIODE

The characteristic curves of a solid-state diode (semiconductor or selenium) may be obtained using the same circuit as that used for the tube provided that corresponding polarity is observed. See Fig. 1-6. Of course, neither the semiconductor nor the selenium diode will require a cathode heater supply. Typical curves for germanium and silicon diodes are shown in Fig. 1-8 on page 8.

1-8. DIFFERENCES BETWEEN TUBE AND SEMICONDUCTOR CHARACTERISTICS

(1) There is evidently some value of "cut-in" or "threshold" voltage before the semiconductor diodes begin to permit current flow, approximately 0.2 V for germanium and 0.6 V for silicon. (A selenium diode exhibits a similar cut-in voltage of about 4.0 V per cell.)

(2) The voltage across the diode where the characteristic curve tends to become linear is much lower than in the tube.

(3) There is no leveling off or saturation of forward current as exists in most thermionic diodes. The current increases exponentially with voltage.

(4) There is a finite reverse or leakage current, I_o, in the solid-state diodes; in the order of microamperes in the germanium diode but only a few nanoamperes in some silicon diodes. Tubes have practically zero reverse current below the PIV.

(5) Although it is not shown in the above curves, the semiconductor diodes are sensitive to the *ambient* temperature (surrounding environment), whereas the tubes are not.

Zener Voltage

One similarity does exist between a tube and a semiconductor diode, and that is the peak inverse voltage restriction. This may be referred to as the breakdown or Zener voltage, V_Z, although only Zener diodes have a sharp breakdown characteristic with a clearly identifiable Zener voltage. Other solid-state diodes have a gradual breakdown as illustrated in Fig. 1-8, and may have a PIV from 50 V to 200 kV depending upon the application. A detailed explanation of the Zener breakdown will be covered in Chapter 6.

1-9. DIODE RESISTANCE

The comparison of various diodes, whether semiconductor or thermionic, can be made quickly by plotting their characteristic curves on the same coordinate axes on graph paper and noting their "steepness" for increasing voltage. This is a measure of their resistance, as explained in Fig. 1-9 for a thermionic diode.

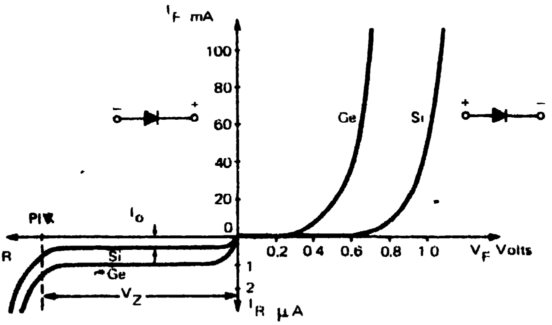


Fig. 1-8. Typical characteristic curves for germanium and silicon diodes. (Note the magnified scale for reverse current.)

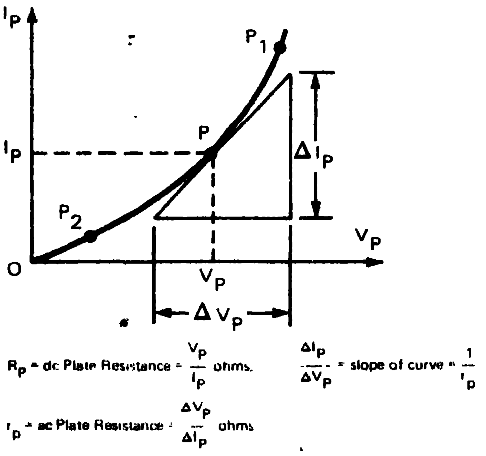

I_P
P_1
P
P_2
I_P
ΔI_P
O
V_P
V_P
ΔV_P
R_P = dc Plate Resistance = $\frac{V_P}{I_P}$ ohms
$\frac{\Delta I_P}{\Delta V_P}$ = slope of curve = $\frac{1}{r_P}$
r_P = ac Plate Resistance = $\frac{\Delta V_P}{\Delta I_P}$ ohms

Fig. 1-9. Curve showing how to calculate the dc and ac resistance of a diode.

R_P is the *static* plate resistance at point P, while r_p is the *dynamic* plate resistance at point P. They will be different from each other and will vary according to the portion of the curve under consideration. At point P_1 for example, the curve is steeper, as can be seen by drawing a tangent to the curve at this point. But since $r_p = \frac{1}{\text{slope}} = \frac{1}{\text{“steepness”}}$ this means r_p is smaller at P_1 than at P. Similarly, r_p is greater at P_2 since the curve is less steep.

Typical values of diode resistance for a tube are as follows: $R_P \approx 500$ ohms; $r_p \approx 100$ ohms; resistance in reverse direction, R_R, in the order of thousands of megohms.

Semiconductor Resistances

The resistance of a semiconductor diode can be determined in the same manner as for the tube shown above, but with a slight difference in notation. The static resistance is referred to as R_F for forward resistance or just R, and the dynamic resistance as R_f or r. Due to the lower voltage drop across the semiconductor diode and also its steeper characteristic curve, it should be evident that both the static and dynamic resistances are lower than for the tube. Typical values are: $R \approx 30$ ohms, $r \approx 1$ ohm; resistance in reverse direction, R_R, in the order of 1 Megohm and less for a germanium diode, but approximately 100 Megohms for a silicon diode.

It may be shown that at room temperature the dynamic resistance of a semiconductor diode having a forward bias greater than its cut-in voltage is given by Shockley's relation:

$$r \approx \frac{26\,n}{I} \qquad (1\text{-}2)$$

Where: n = 1 for a germanium diode and n = 2 for a silicon diode; r is in ohms if I is the dc current in milliamperes.*

The dynamic resistance is a very important parameter and although it varies with current, it may be considered constant for small voltage variations.

Meaning of Static and Dynamic Resistance

The static resistance R is the resistance offered by the device when only a steady dc current is flowing through the diode, such as I_P at P in Fig. 1-9. If the current through the device is varying above and below some average value of current (such as I_P), the diode has an effective ac resistance called the dynamic resistance r, to the varying component of the current i_p. The larger the average value of current flowing through the diode the smaller will be its dynamic resistance, r, since the operating or quiescent point is on a steeper portion of the curve.

1-10. COMPARISON OF PRACTICAL DIODES WITH THE IDEAL DIODE

It should be fairly evident from a comparison of the characteristic curves of the semiconductor diodes and the thermionic diode which one more closely approaches the ideal diode's characteristics. When the compactness of the silicon diode is also considered it seems to have an overwhelming superiority compared with the tube. Perhaps it is fitting that the symbol used to represent the ideal diode is the same as that of the semiconductor diode. However, if very high voltage and power are involved, if practically zero reverse current is required, or if the diode is to be located in an environment of high temperature and nuclear radiation, then a tube or selenium diode will continue to be specified.

1-11. EQUIVALENT CIRCUITS

The analysis of a circuit containing a diode may be carried out in a number of ways. One is a graphical method and uses the actual characteristic curve of the diode. This gives a more accurate result and will be covered in the next chapter. However, for many circuits it is sufficient to use a "straight-line approximation" for the non-linear device; the process of obtaining such an approximation is called "linearizing."

A practical diode's characteristic curve can be linearized in either of two ways as shown in Fig. 1-10 with their equivalent circuits, or models.

In Fig. 1-10(a) the curve has been approximated by a straight line through the origin and the best average position of the rest of the curve. The plate resistance R_P corresponds to the static plate resistance at the intersection. In this equivalent diagram, whenever P is more positive than K, the ideal diode will switch on, and the only effect will be that of a resistance of value R_P in any given circuit.

In Fig. 1-10(b) a closer approximation to the curve has been made, with the assumption that the diode will not turn on until the voltage at P with respect to K exceeds the voltage represented by the battery, V_o. The plate resistance, r_p, now corresponds to the dynamic resistance where the line was drawn tangent to the curve, but this is not meant to imply that this equivalent circuit is good only for ac; it is valid for both ac and dc applications.

The equivalent circuit utilizing the battery, V_o, is eminently suited for representing the semiconductor diodes, since V_o can represent the cut-in voltage of 0.2 V for germanium or 0.6 V for silicon. Also, r represents the average dynamic resistance of the diode in the almost linear portion of the curve, and is

---

*Millman, J., and Halkias, C., *Integrated Electronics*, McGraw-Hill, New York, 1972, Section 3-6.

the value most significant in rectifying applications.

However, even the thermionic diode can be represented by the battery equivalent circuit since it is seldom (except in radio receiver detection circuits) that the device is operated in the region of extreme curvature near the origin, and r_p is much more nearly constant than is R_P.

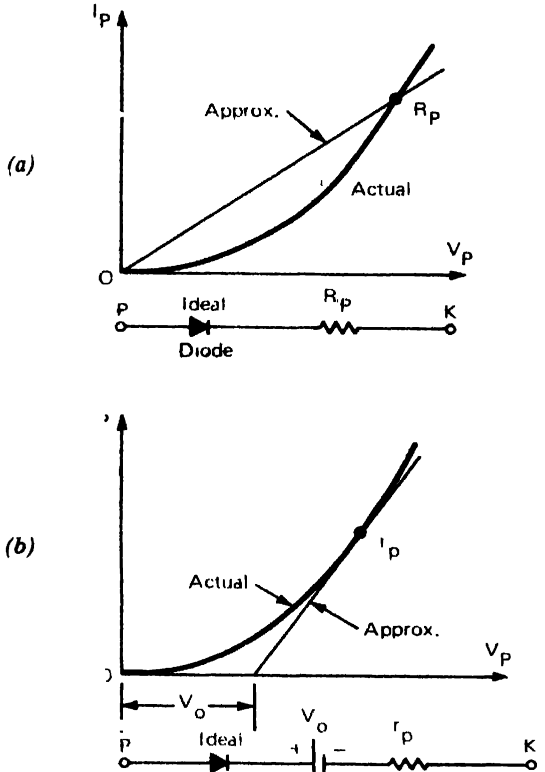


Fig. 1-10. Methods of linearizing a diode's characteristics. (a) The straight line is drawn through the origin. (b) A cut-in voltage is approximated to make a closer approximation to the curve in the portion of interest.

The equivalent circuits do not allow for the small reverse currents that flow in solid-state devices up to the PIV rating, or Zener voltage. Modification of the diagram can be made if desired to take this into account, but except for critical applications and cases where the Zener voltage is employed to advantage this reverse current is ignored.

1-12. ANALYSIS OF A CIRCUIT CONTAINING A DIODE

Let us consider an example of the use of an equivalent circuit to analyze a network containing a diode.

Example 1-2. Determine the waveform of the output voltage, v_o, in the circuit of Fig. 1-11(a), and the waveform of the voltage across the diode, v_D. The diode's characteristics are represented by the equivalent circuit of Fig. 1-11(b), and the input voltage, v_s, has the waveform shown in Fig. 1-12(a).

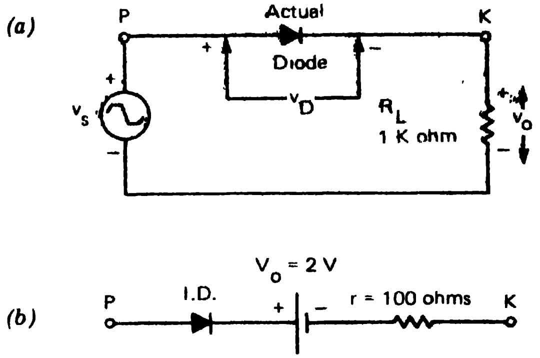


Fig. 1-11. (a) Schematic of circuit to be analyzed for v_O and v_D. (b) Equivalent circuit of the actual diode.

Solution. Assuming for a moment the diode has no resistance, (r = 0), the peak of the output voltage would be:

$$V_{sm} - V_o = 10 - 2 = 8 \text{ V},$$

because the diode will not conduct until the input voltage exceeds 2 V, as shown in Fig. 1-12(b).

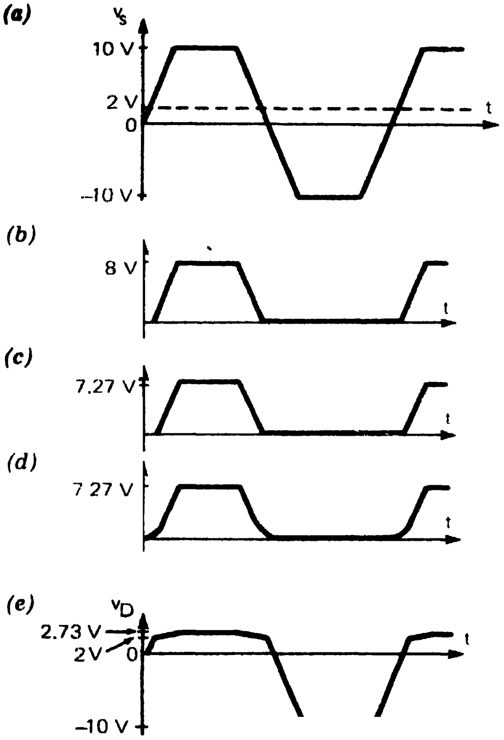


Fig. 1-12. Voltage waveforms in the rectifying circuit of Fig. 1-11.
(a) Input voltage v_s, showing the 2 V, V_o level of the diode's equivalent circuit.
(b) Output voltage v_o, assuming no diode forward resistance.
(c) Output voltage v_o, taking into account the diode's average dynamic resistance.
(d) Output voltage v_o, showing the distortion that would occur if the diode's actual characteristics had been used.
(e) Diode voltage v_D, which is the difference between the waveforms in (a) and (c).

But the effect of the resistance r is to cause an additional voltage drop across the diode, resulting in a decreased output voltage. The maximum output voltage is now given by the voltage divider equation:

$$8\text{ V} \times \frac{R_L}{r + R_L} = 8\text{ V} \times \frac{1000}{100 + 1000}$$

$$= 8\text{ V} \times \frac{10}{11} = 7.27\text{ volts},$$

as shown in Fig. 1-12(c).

Had the actual characteristics of the diode been used, the output would be distorted at low voltage levels due to the curvature of the characteristic near the origin, as illustrated by Fig. 1-12(d).

Fig. 1-12(e) shows the waveform of voltage that will appear across the diode using the equivalent circuit of the diode. It is clear that the forward voltage drop across the diode is not negligible in this example.

Note that the waveform of v_D may be obtained from Kirchoff's voltage law:

$$v_s = v_D + v_o \tag{1-3}$$

$$\therefore v_D = v_s - v_o \tag{1-4}$$

and is drawn by subtracting the curve in Fig. 1-12(c) from that in Fig. 1-12(a). The maximum reverse voltage that occurs across the diode is 10 V, so that the PIV rating of the diode must be higher than this value.

The output v_o is said to be a half-wave rectified version of the input and is called pulsating dc, since the negative portion of the input has been removed. This is the primary purpose of diodes and will be examined very fully in the next chapter on dc power supplies.

1-13. TUBE TYPE NUMBERS

Although the numbering system of tubes does have more meaning than the numbers assigned to transistors it is still necessary to become acquainted with the functions of the tube through experience or by referring to tube manuals available from the leading manufacturers such as RCA, GE, etc.

However, some general rules do apply to tubes. Consider a 5Y3-GT. This is a tube containing two plates and one common directly heated cathode. See Fig. 1-13.

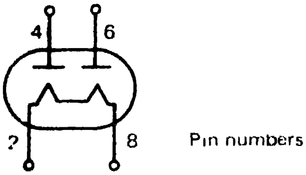


Fig. 1-13. Symbol of a duodiode containing two plates and a directly heated cathode.

It is thus known as a duodiode and is used widely in full-wave rectifiers. The "5" in 5Y3-GT indicates 5 volts are necessary at the filament for correct electron emission. The "3" indicates there are three elements connected to the base by means of four connecting pins. To determine which pin is which requires a tube manual.

1-14. SEMICONDUCTOR NUMBERS

Semiconductor diodes have a "1" prefix such as 1N2615 whereas transistors have a "2" prefix, such as 2N217, or just a five-digit number, such as 40340. However, no distinction is made in the numbering system between silicon and germanium devices for either diodes or transistors. One may generalize and say that the larger the number of the device the more recent its development, but it is a pity and inconvenience that no distinction is made between the various materials of the semiconductor devices. A manual is the only recourse to determine what a number represents. Semiconductor diodes may be bullet shaped (similar to the symbol), have a diode symbol printed on the body to show the direction of forward current flow, or have a band on the body to indicate the "cathode." However, if no polarity marking is available, a quick check with an ohmmeter (whose leads are known for plus and minus) will soon reveal the normal forward direction.

1-15. SUMMARY

(1) The thermionic vacuum diode has a plate and either a directly or an indirectly heated cathode. It is bulky when compared with the selenium diode or the ultra compact semiconductor diode.

(2) An ideal diode is a two terminal polarity sensitive device that has zero resistance when it is forward biased and infinite resistance when reverse biased.

(3) A vacuum diode's characteristics follow Child's three halves power law, $I = KV^{3/2}$, while current increases exponentially with voltage in a semiconductor diode.

(4) Silicon and germanium diodes exhibit a cut-in voltage of 0.6 V and 0.2 V respectively in their characteristic curves, and thus approximate closely the ideal diode in this respect.

(5) The peak inverse voltage, PIV, is the highest reverse voltage a diode can withstand before breaking down and permitting current to flow in the reverse direction.

(6) The static and dynamic resistances of a diode may be determined from the characteristic curves and represent the dc and ac opposition to current flow offered by the diode.

(7) An equivalent circuit of a diode is a linearization of the diode's characteristic curve and typically consists of an ideal diode in series with a battery (representing the cut-in voltage), and a resistance equal to the dynamic resistance of the diode over its linear portion.

(8) A diode in a series circuit with an ac voltage applied will cause half-wave rectification.

QUESTIONS

Q1-1. In what type of equipment are directly heated cathode tubes used?

Q1-2. What is a duodiode? Draw its symbol with an indirectly heated cathode.

Q1-3. What are the three main types of diodes?

Q1-4. What are the approximate volumes, in cubic inches, occupied by a 5Y3-GT duodiode, a selenium diode @ 75 mA and 500 V, and a typical silicon diode @ 1 A? (See references to these in the text.)

Q1-5. What resistance does an *ideal* diode have when it is on, and when it is off?

Q1-6. What does the + on a selenium diode indicate? Does this correspond to the voltage drop across the diode?

Q1-7. What do you understand by the terms "forward-biased" and "reverse-biased" diodes?

Q1-8. Explain in your own words the meaning of the following terms:

(a) virtual cathode,
(b) plate current saturation, as applied to a vacuum diode.

Q1-9. What is meant by the PIV of a diode? Does this term apply equally well to all diodes? What alternative terms are used for silicon diodes?

Q1-10. What does Child's Law state? Does this apply only to tubes?

Q1-11. What is a mathematical model?

Q1-12. Sketch the V-I characteristics of typical silicon and germanium diodes and indicate the approximate "cut-in" or "threshold" voltages for each.

Q1-13. Give at least two advantages and two disadvantages of the semiconductor diodes compared with the tube diode.

Q1-14. As current through a vacuum diode is increased what happens to the static resistance? to the dynamic resistance?

Q1-15. *Sketch* a curve showing how the dynamic resistance of a silicon diode varies with voltage across the diode, from 0 V to 1 V, without using actual values for resistance. Indicate the cut-in voltage and explain its significance in determining diode resistance.

Q1-16. Explain what you understand by the expression, "linearizing a non-linear device."

Q1-17. What effect does the cut-in voltage of a silicon diode have on the output waveform across a resistor when placed in series with a diode and a sinusoidal input voltage of 2 V peak-to-peak? What is the name of this output?

Q1-18. A silicon diode is being checked by a student with an ohmmeter. What will be the relative values indicated as he switches from the X1 range to X10, X100, X1K, X100K, assuming the diode is forward biased?

PROBLEMS

P1-1. A selenium diode is to be used in a 117 V half-wave rectifier. What is the minimum number of cells that will be needed to be stacked in series to provide the required PIV capability? How many would you suggest to provide some margin of safety? (Line voltages may go as high as 125 V.)

P1-2. What would be the minimum forward-voltage drops across the selenium rectifiers in P1-1?

P1-3. A vacuum diode has a forward-voltage drop of 16 V when passing a current of 50 mA. It is known to be operating in its space-charge-limited region. Determine:

(a) An equation relating I_P to V_P.
(b) The plate current expected when the voltage drop is 25 V.
(c) The voltage across the diode when the current is 25 mA.

P1-4. A 6AL5 tube passes 15 mA with 4 V drop across it.

(a) How much current will flow with 8 V across the tube?
(b) What voltage drop will exist when 7.5 mA is passing through the tube?

P1-5. A student finds experimentally that a tube has a Child's Law K = 0.2 (current in milliamperes, voltage in volts). What voltage drop can he expect across this tube at a current of 12.8 mA?

P1-6. What is the static or dc resistance of the following:

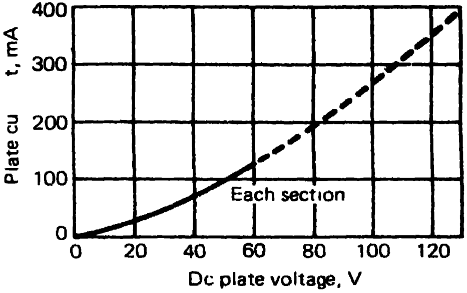
(a) A 6AL5 operating at 4 V and 15 mA
(b) A 5Y3 operating at 40 V and 75 mA

P1-7. Referring to the information in P1-6 determine the ratio of the current in the 5Y3 to the current in the 6AL5 if the two tubes are operated in parallel, and both obey Child's Law.

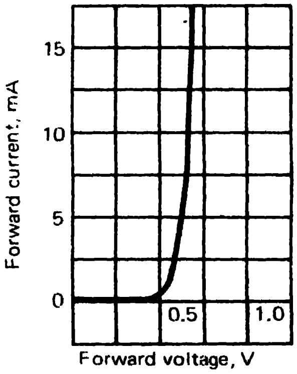
P1-8. *Plot* graphs on the same axes, showing how the dynamic resistances of silicon and germanium diodes vary with current flowing through the diodes, using Eq. 1-2. What is the name of such curves?

P1-9. Using the characteristic curve of the 5Y3-GT diode obtain an equivalent linear circuit by: (a) Drawing a straight line through the origin and the best average position of the curve over the range of current up to 125 mA. Determine R_P. (b) Drawing a tangent to the curve at 75 mA, and determining V_o and r_p. Draw the two equivalent circuits.

Characteristic curve of 5Y3-GT diode for P1-9

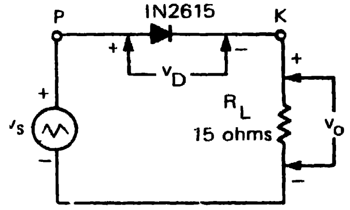


P1-10. Obtain an equivalent linear circuit, utilizing V_o and r, for the 1N2615 silicon diode whose characteristics are given.

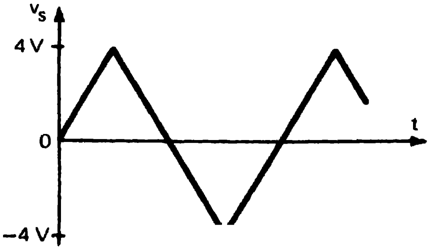


1N2615 Characteristic curve for P1-10.

P1-11. Using the equivalent circuit of the silicon diode from P1-10 determine the waveforms of the output voltage v_o and the diode voltage v_D in the circuit shown with the given input waveform.



P1-11

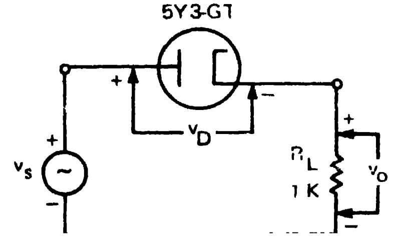


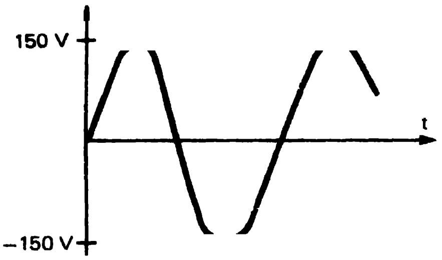
P1-12. Repeat problem P1-11 using an 8 V p-p sinusoidal input.

P1-13. Repeat problem P1-11 using an 8 V p-p square-wave input.

P1-14. A half-wave rectifier circuit using a 5Y3-GT diode is in series with a 1 K load as shown. For the given input determine the waveforms of v_o and v_D using:

(a) The equivalent circuit obtained in P1-9(a).
(b) The equivalent circuit obtained in P1-9(b).

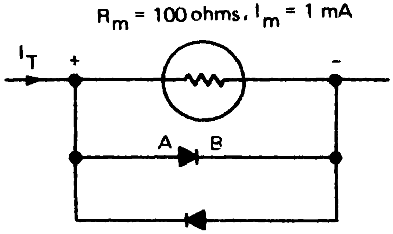


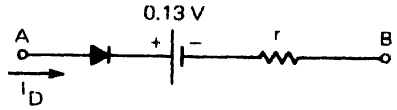


P1-15. What must be the minimum PIV rating of the diode in P1-14?

P1-16. Repeat P1-14 using a 50 V p-p square-wave input.

P1-17. A common way of protecting a movement in a dc milliammeter from an overload of current in either direction is to connect two germanium diodes across it as shown. These provide a shunting effect when I_T becomes excessive.





Given the above movement of resistance 100 ohms and 1 mA full scale deflection, let each diode be represented by the given equivalent circuit.

The resistance of the diode, r, varies with the current I_D through the diode according to the equation $r = 26/I_D$ where r is in ohms when I_D is in mA.

(a) When I_T = 1 mA how much current flows through the movement?
(b) When I_T = 5 mA how much current flows through the movement?
(c) When I_T = 10 mA in the opposite direction how much current will flow through the movement?

2

RECTIFYING CIRCUITS AND DC POWER SUPPLIES

2-0. INTRODUCTION

Most electronic circuits need a source of dc power. In some cases a simple power supply is sufficient where the line voltage is converted to dc through half-wave or full-wave rectification. This is called unidirectional pulsating current. A later chapter will cover the subject of filtering this current to provide smooth dc.

In this chapter, the method of analyzing a diode circuit, using a load line drawn on the characteristic curve of the diode, is introduced. This is a very powerful tool and one which will be used extensively with other active devices. Half-wave and full-wave rectifying circuits are examined and compared, and a method is described to determine the size of transformer that should be used to deliver a given amount of dc power.

2-1. LOAD LINE ANALYSIS OF A DIODE CIRCUIT

If the equivalent linear circuit of a diode does not provide an accurate enough result in some applications, the actual characteristic curve of the diode must be used for analysis.

DC Applied Voltage

Consider the determination of the current that will flow in the circuit of Fig. 2-1(a) and what the output voltage V_o will be, given the characteristic curve shown in Fig. 2-1(b).

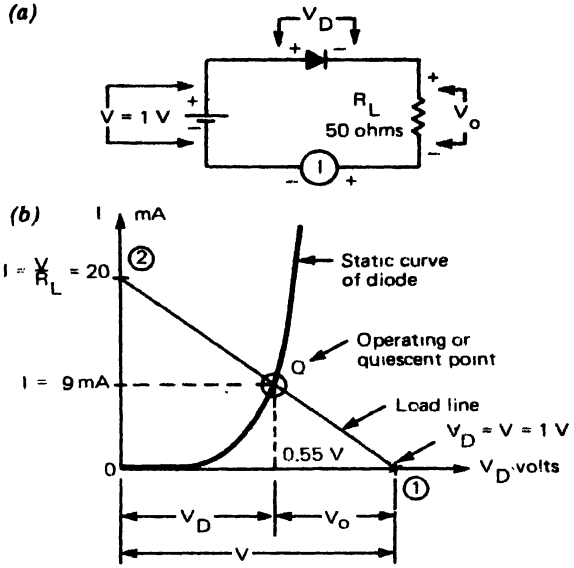


Fig. 2-1. (a) Diode circuit to be analyzed for I and V_o. (b) Characteristic curve with load line.

How much current will flow in the circuit will depend upon the static resistance of the diode. However, the resistance of the diode depends upon how much current flows in the circuit! Where is one to start? Consider Kirchoff's voltage law applied to the circuit:

$$V = V_D + V_o = V_D + IR_L. \qquad (2\text{-}1)$$

This is an equation that relates the two variables involved, I and V_D, and represents a number of pairs of coordinates for a given V and R_L. Let us consider two such pairs:

(1) Consider $I = 0$; then $V = V_D = 1$ V. This gives one point, (1,0), plotted on the horizontal axis.

(2) Consider $V_D = 0$; then $I = \frac{V}{R_L} = \frac{1}{50}$ A = 20 mA.

This is another point, (0,20), plotted on the vertical axis.

A line drawn to connect these two points represents all the possible combinations of I and V_D that satisfy our initial equation.

The intersection of this line, the load line, with the static characteristic curve of the diode provides the operating point, sometimes called the quiescent point or Q point. This is a point of equilibrium at which just enough current flows in the circuit to cause the correct voltage drop to occur across the diode so that the output voltage is the balance of the input voltage.

In this example the current in the circuit is 9 mA, the voltage drop across the diode is 0.55 V, and V_o is 0.45 V.

The static resistance of the diode is given by:

$$R = \frac{V_D}{I} = \frac{0.55}{9 \times 10^{-3}} \text{ ohms} = 61 \text{ ohms}$$

and the total resistance in the circuit by:

$$R + R_L = 61 + 50 = 111 \text{ ohms.}$$

That amount of resistance will allow 9 mA of current to flow when 1 V is applied, which checks our solution.

Equation of the Load Line

The line drawn between the two points found above is called the load line since its slope depends upon the value of the load as shown below:

$$\text{Given:} \quad V = V_D + IR_L \tag{2-2}$$

$$\text{Then:} \quad IR_L = V - V_D = -V_D + V \tag{2-3}$$

$$\text{and:} \quad I = -\frac{V_D}{R_L} + \frac{V}{R_L} \tag{2-4}$$

$$\text{or:} \quad I = -\frac{1}{R_L} V_D + \frac{V}{R_L}. \tag{2-5}$$

If this is compared with the standard equation of a straight line, of y vs. x, having a slope m and y intercept c,

$$y = m x + c \tag{2-6}$$

it can be seen that the load line has a slope given by $-1/R_L$ and an intercept on the I axis (y axis) of V/R_L. Thus for a given input voltage V, the effect of a higher load (smaller value of R_L) will be a steeper load line and a resulting higher value of current at the Q point. In a later chapter it will be necessary to draw additional load lines through given Q points, with slopes equal to $-1/R_L$. This will not be pursued here, but it should be pointed out that the minus sign indicates a line that slopes from the upper left to the lower right, on the characteristic curve.

AC Applied Voltage

Now consider the application of an alternating voltage to the circuit of Fig. 2-1(a), as shown in Fig. 2-2(a). Let v_s have a 2 V peak-to-peak sinusoidal alternation and let us determine the waveform of output current (and voltage) using the same characteristic curve for the diode pictured in Fig. 2-2(b).

Since the voltage applied to the diode-resistor combination is varying, the process of drawing a load line for the instantaneous values of v_s must be repeated at convenient intervals. However, since the load R_L is constant, all these load lines will be parallel. The intersection of these load lines with the static curve of the diode will give the current in the circuit corresponding to each instantaneous value of input voltage.

Instead of using these values of current to obtain the output current waveform, it is more convenient to construct a dynamic curve of i versus v_s as shown in Fig. 2-2(b).

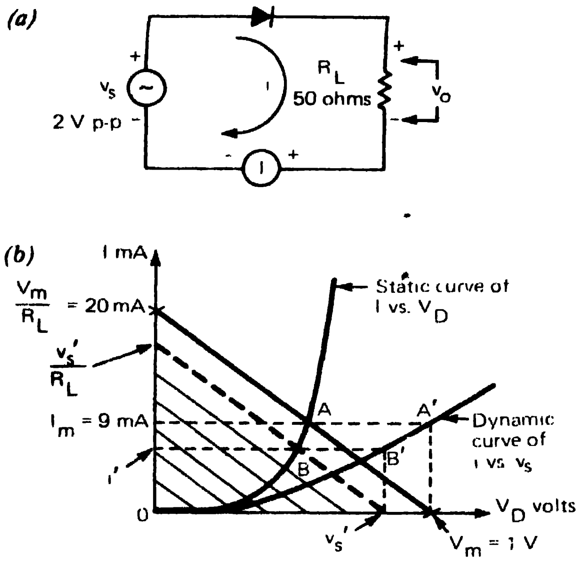


Fig. 2-2. (a) Diode circuit to be analyzed for waveforms of i and v_O. (b) Showing how to determine the dynamic curve of i vs. v_s for a given R_L from the static curve of the diode.

Consider when v_s is at its maximum, V_m = 1 V. The intersection of the load line corresponding to this condition with the static curve is shown to occur at A, giving a peak current in the circuit of I_m = 9 mA. Drawing a line horizontally from A will intersect a vertical line drawn from V_m at A'. Now take any other value of v_s, say v_s' and draw a load line parallel to the first. The intersection at B gives the current in the circuit, i', corresponding to v_s'. Drawing a line horizontally from B and vertically from v_s' gives another point B'. Continuing this process for arbitrary values of v_s will yield sufficient points to construct the dynamic curve shown in Fig. 2-2(b).

The horizontal axis in this graph must now be referred to as v_s when using the dynamic curve. It must be emphasized that this dynamic curve applies only to the circuit containing the value of load resistance for which it was drawn.

The manner in which the output current waveform is obtained using the diode's dynamic curve (drawn for R_L = 50 ohms) is indicated graphically in Fig. 2-3. The input voltage v_s is drawn with its time axis vertically so that values of voltage for different time intervals, t_1, t_2, etc., may be projected onto the voltage axis of the diode's dynamic curve. Intersections with the curve are drawn horizontally to give the values of current corresponding to the times t_1, t_2, etc.

It is convenient if the time axis of the output current graph corresponds in scale to the input voltage time axis, but it is not essential. In Fig. 2-3, the times chosen, t_1, t_2, and t_3, correspond to 30°, 60°, and 90° of the input waveform. These usually give sufficient points to sketch the output quite accurately.

It is obvious that the output current is distorted on the positive half cycle. Look at the values of voltage and current at t_1. Whereas v_s is one half of its maximum value, 0.5 V, the output current is only about one-quarter of its maximum value at t_1. In addition, due to the cut-in voltage of most semiconductor diodes, the output current does not appear until some time after the input voltage has been applied. Consequently the period of conduction of the diode is less than half a cycle. For values of input voltage greatly in excess of the cut-in voltage the portion of the dynamic curve that determines the output current (and voltage) will be practically linear, with the result that output waveforms will be substantially sinusoidal over the positive half cycles. Assuming that they are sinusoidal will allow us to make some relatively simple conclusions regarding the readings of voltmeters and ammeters when connected in rectifying circuits, as shown in the next section. But where small signals are applied to diode circuits, such as in communication detector circuits, the nonlinearity in output current is of prime importance, and the above discussion very valuable.

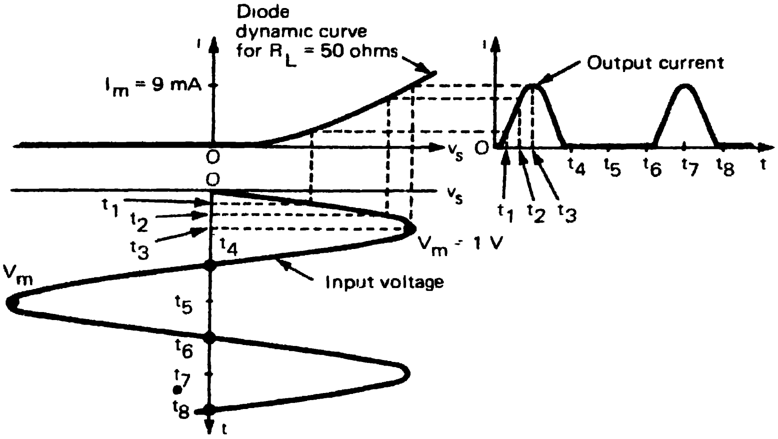

Diode dynamic curve for R_L = 50 ohms
i
I_m = 9 mA
O
v_s
O
v_s
t_1
t_2
t_3
t_4
V_m = 1 V
V_m
Input voltage
t_5
t_6
t_7
t_8
t
i
Output current
O
t_1 t_2 t_3
t_4 t_5 t_6 t_7 t_8 t

Fig. 2-3. Showing how to determine the waveform of output current (and voltage) for a given input v_s using the diode's dynamic curve which was drawn for R_L = 50 ohms.

2-2. THE HALF-WAVE RECTIFIER

In the previous section we determined the waveform of current, but did not consider what a dc ammeter would indicate in a half-wave rectifying circuit, or what the dc output voltage would be as indicated by a dc voltmeter. For this discussion refer to Fig. 2-4.

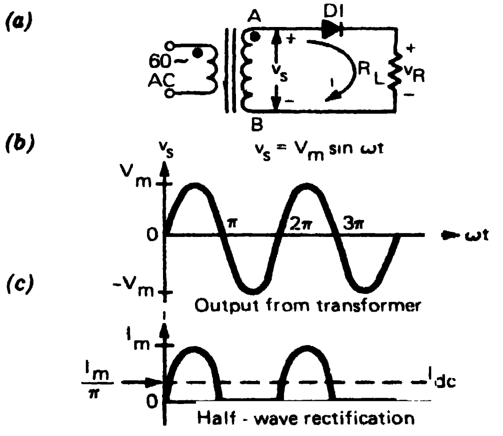

(a)
A
DI
60~ AC
v_s
+
-
i
R_L
v_R
+
-
B
(b)
v_s
$v_s = V_m \sin \omega t$
V_m
0
π
2π
3π
ωt
(c)
$-V_m$
Output from transformer
i
I_m
$\frac{I_m}{\pi}$
0
I_{dc}
Half - wave rectification

Fig. 2-4. (a) Half-wave rectifying circuit. (b) Open circuit voltage from transformer. (c) Load current showing the average value, I_{dc}.

A transformer has been added to provide the desired voltage into the rectifier. For high voltage dc power supplies this would obviously be a step-up transformer, but with many solid-state equipment applications voltages of 4.5, 6, 9, 22.5, 40 volts etc. are used, thus necessitating a step-down transformer.

DC Output Current

Assuming the open circuit voltage from the secondary of the transformer is given by

$$v_s = V_m \sin \omega t \tag{2-7}$$

the peak value of load (and diode) current will be:

$$I_m = \frac{V_m}{R_s + R_f + R_L} \tag{2-8}$$

where R_s is the resistance of the transformer secondary

R_f is the dynamic resistance of the diode

R_L is the load resistance.

This assumes the cut-in voltage of the diode is small enough compared with V_m to be neglected (not always the case in low voltage power supplies) and allows us to consider the current i to consist of a half-wave rectified sinusoid. It is evident that a dc ammeter in the circuit will not follow the variations of i but will assume some average position. But what is the average of i over an interval from 0 to 2π ? This may be answered by defining the average to be the area under the curve from 0 to π divided by the base, 2π, and may be obtained mathematically by using calculus. Calling this average value I_{dc} the result is:

$$I_{dc} = \frac{I_m}{\pi}. \qquad (2\text{-}9)$$

This may also be seen from the following equation for i, which represents a Fourier series analysis* of the output current:

$$i \approx \frac{I_m}{\pi} + \frac{I_m}{2}\sin\omega t - \frac{2\,I_m}{3\pi}\cos 2\,\omega t - \frac{2\,I_m}{15\pi}\cos 4\,\omega t. \qquad (2\text{-}10)$$

Only the first four terms in the series have been indicated. It is evident that the output current consists of a dc component, I_m/π, and alternating components whose average values are zero, and thus contribute nothing to the dc ammeter reading.

DC Output Voltage

The dc output voltage appearing across R_L will be:

$$V_{dc} = I_{dc}\,R_L \qquad (2\text{-}11)$$

which can be put in terms of V_m by substituting the expressions for I_{dc} and I_m.

$$V_{dc} = I_{dc}\,R_L \qquad (2\text{-}12)$$

$$= \frac{I_m}{\pi}\,R_L \qquad (2\text{-}13)$$

$$= \frac{\frac{V_m}{\pi}\,R_L}{R_s + R_f + R_L}, \qquad (2\text{-}14)$$

$$= \frac{\frac{V_m}{\pi}}{\frac{R_s + R_f}{R_L} + 1}. \qquad (2\text{-}15)$$

Thus if $R_L >> R_s + R_f$ the output voltage is:

$$V_{dc} \approx \frac{V_m}{\pi}. \qquad (2\text{-}16)$$

Alternatively,

$$I_{dc} = \frac{I_m}{\pi} \qquad (2\text{-}17)$$

$$= \frac{\frac{V_m}{\pi}}{R_s + R_f + R_L} \qquad (2\text{-}18)$$

from which, solving for $V_{dc} = I_{dc}\,R_L$, we obtain

$$V_{dc} = \frac{V_m}{\pi} - I_{dc}\,(R_s + R_f). \qquad (2\text{-}19)$$

This analysis clearly indicates that the dc output voltage depends not only upon the load current but the diode and transformer resistances also.

2-3. VOLTAGE REGULATION

The degree to which a power supply varies in output voltage under conditions of load variations is measured by the voltage regulation, usually expressed as a percentage:

---

*A mathematical method of expressing a complex waveform as a sum of dc and ac components.

% Voltage Regulation =

$$\frac{V_{\text{no-load}} - V_{\text{full-load}}}{V_{\text{full-load}}} \times 100\%. \qquad (2\text{-}20)$$

A Thevenin equivalent circuit for the half-wave rectifier is shown in Fig. 2-5, where R_{INT} is the internal resistance of the power supply and consists in this case of $R_s + R_f$.

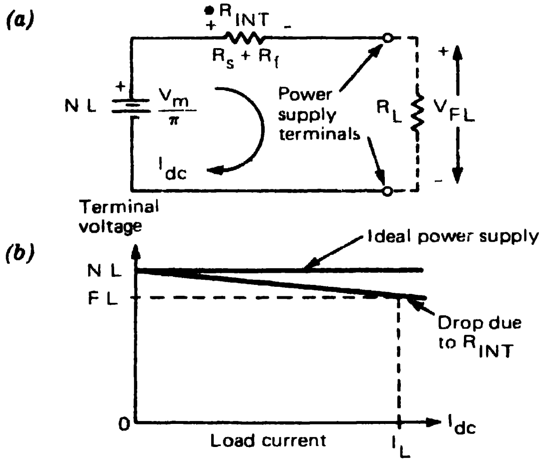


Fig. 2-5. (a) Thevenin equivalent circuit of a half-wave rectifier dc power supply. (b) Graph showing variation in terminal voltage with load current, for an ideal power supply and a practical power supply.

Note that a power supply that has good voltage regulation has a low percentage regulation. An ideal power supply has a full-load voltage equal to its no-load voltage, hence zero percent regulation.

If, for example, a power supply had a no-load voltage of 9 V and a full-load voltage of 8 V, the percentage regulation would be:

$$\frac{9-8}{8} \times 100\% = 12.5\%.$$

The value of the internal resistance can be determined experimentally as follows: Record the no-load (open-circuit) voltage of the power supply with a high resistance voltmeter. Then load the power supply, noting the full-load voltage and current. The difference between no-load and full-load voltage is the amount of voltage drop that occurred across the internal resistance of the power supply.

That is,

$$V_{NL} - V_{FL} = I_L R_{INT}. \qquad (2\text{-}21)$$

Hence:

$$R_{INT} = \frac{V_{NL} - V_{FL}}{I_L} \text{ ohms} \qquad (2\text{-}22)$$

where I_L is in amperes.

By using suitable voltage regulators involving regulator tubes, Zener diodes and transistors, the internal resistance can be reduced to a fraction of an ohm. Such regulators will be considered later.

2-4. RIPPLE FACTOR

The object of a rectifier is to convert the ac to dc. A measure of how successful a circuit is in doing this is the ripple factor, r. This is defined as follows:

$$r = \frac{\text{ripple voltage}}{\text{dc voltage}}$$

$$= \frac{\text{rms value of ac components}}{\text{dc value of wave}}$$

$$= \frac{V_{r,rms}}{V_{dc}} \qquad (2\text{-}23)$$

where $V_{r,rms}$ is the rms value of the ripple voltage. The ripple factor may also be defined in terms of ripple current and dc current:

$$r = \frac{I_{r,rms}}{I_{dc}}. \qquad (2\text{-}24)$$

Let us calculate the ripple factor of the half-wave rectifier. Recall that the expression for current in the half-wave rectifier is given by:

$$i \approx \frac{I_m}{} \sin \omega t - \frac{2\,I_m}{3\pi} \cos 2\,\omega t \quad (2\text{-}25)$$

$$\frac{2\,I_m}{15\pi} \cos 4\,\omega t$$

$$\text{or } i = I_{dc} + i_r \qquad (2\text{-}26)$$

where $I_{dc} =$

$$\text{and } i_r = \frac{I_m}{2} \sin \omega t - \frac{I_m}{\pi} \cos 2\,\omega t$$

$$\frac{2\,I_m}{15\pi} \cos 4\,\omega t$$

Let I_{rms} be the rms value of i, the total half-wave rectified output current. Similarly, let $I_{r,rms}$ be the rms value of the ripple current, i_r, consisting of the alternating components. Then, from ac circuit theory:

$$I_{rms}{}^2 = I_{dc}{}^2 + I_{r,rms}{}^2 \qquad (2\text{-}27)$$

and

$$I_{r,rms} = \sqrt{I_{rms}{}^2 - I_{dc}{}^2}. \qquad (2\text{-}28)$$

$$\text{Thus, } r = \frac{I_{r,rms}}{I_{dc}} \qquad (2\text{-}29)$$

$$= \frac{\sqrt{I_{rms}{}^2 - I_{dc}{}^2}}{I_{dc}} \qquad (2\text{-}30)$$

$$= \sqrt{\left(\frac{I_{rms}}{}\right)^2} \qquad (2\text{-}31)$$

It is now required to know the value of I_{rms}. That is, what an ac ammeter that responds to *true rms* current would indicate when placed in the half-wave circuit. (Many ac ammeters convert the ac to dc and use a D'Arsonval movement for the indication. Such a meter will not indicate true rms current. A "square-law" instrument such as a thermocouple or moving iron vane type is necessary.)

The rms value of i, I_{rms}, may be found by using calculus to be:

$$I_{rms} = \frac{I_m}{2}. \qquad (2\text{-}32)$$

Note how this is different from the rms value of a sinusoidal wave, $I_m/\sqrt{2}$. The value for I_{rms} may also be found from the expression for i by finding the square root of the sum of the squares of the rms values of each of the components in i. (See Problem 2-5.)

Thus,

$$r = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} \qquad (2\text{-}33)$$

$$= \sqrt{\left(\frac{\frac{I_m}{2}}{\frac{I_m}{\pi}}\right)^2 - 1} \qquad (2\text{-}34)$$

$$= \sqrt{\frac{\pi^2}{4} - 1} = 1.21. \qquad (2\text{-}35)$$

Expressed as a percentage, this indicates that the amount of ac present in the output is 121% of the dc voltage. The half-wave rectifie is not very successful in converting ac to dc!

2-5. RATIO OF RECTIFICATION

Another figure used as a measure of merit to compare rectifiers is the ratio of rectification, defined as: Ratio of Rectification

$$= \frac{\text{dc power delivered to the load}}{\text{ac input power from transformer secondary}}$$

$$= \frac{P_{dc}}{P_{ac}} \tag{2-36}$$

This ratio can actually be obtained from a knowledge of the ripple factor (see Problem 2-12), but it is instructive to find it using the definition.

$$P_{dc} = I_{dc}^{2} R_L \tag{2-37}$$

$$= \left(\frac{I_m}{\pi}\right)^2 R_L \tag{2-38}$$

$$= \frac{I_m^{2}}{\pi^2} R_L \tag{2-39}$$

P_{ac} is what a wattmeter would indicate if placed in the half-wave rectifying circuit with its voltage terminals connected across the transformer secondary.

$$P_{ac} = I_{rms}^{2}(R_f + R_L) \tag{2-40}$$

$$= \left(\frac{I_m}{2}\right)^2 (R_f + R_L) \tag{2-41}$$

$$= \frac{I_m^{2}}{4} (R_f + R_L) \tag{2-42}$$

$$\text{Ratio of Rectification} = \frac{P_{dc}}{P_{ac}}$$

$$= \frac{\dfrac{I_m^{2}}{\pi^2} R_L}{\dfrac{I_m^{2}}{4} (R_f + R_L)} \tag{2-43}$$

$$= \frac{4}{\pi^2} \frac{R_L}{R_f + R_L} \tag{2-44}$$

$$= \frac{4}{\pi^2} \frac{1}{\dfrac{R_f}{R_L} + 1} \tag{2-45}$$

If we assume $R_f << R_L$, then

$$\text{Ratio of Rectification} = \frac{4}{\pi^2} = 0.406. \tag{2-46}$$

This means that under the best conditions (no diode loss), only 40.6% of the ac input power is converted into dc power in the load. The rest exists as ac power in the load.

2-6. TRANSFORMER UTILIZATION FACTOR

In any power supply design it is necessary to determine the rating of the transformer. This is done with a knowledge of the dc power to be delivered to the load and the type of rectifying circuit to be used. Let us determine the T.U.F. (transformer utilization factor) for a half-wave rectifying circuit.

$$\text{T.U.F.} = \frac{\text{dc power delivered to the load}}{\text{ac rating of the transformer secondary}}$$

$$= \frac{P_{dc}}{P_{ac}\ \text{rated}} \tag{2-47}$$

At first sight this may seem to be the same as for the ratio of rectification, but the rating of the transformer secondary is different from the actual power delivered by the secondary.

The *rated* voltage of the secondary is V or $V_m/\sqrt{2}$, but the actual rms current flowing through the winding is only $I_m/2$, not $I_m/\sqrt{2}$.

$$\text{Thus, T.U.F.} = \frac{P_{dc}}{P_{ac}\ \text{rated}} = \frac{\dfrac{I_m^{2}}{\pi^2} R_L}{\dfrac{V_m}{\sqrt{2}} \times \dfrac{I_m}{2}} \tag{2-48}$$

but $V_m = I_m(R_f + R_L)$

$$\therefore \text{T.U.F.} = \frac{\dfrac{I_m{}^2}{\pi^2} R_L}{\dfrac{I_m(R_f + R_L)}{\sqrt{2}} \times \dfrac{I_m}{2}} \tag{2-49}$$

$$= \frac{2\sqrt{2}}{\pi^2} \quad \frac{R_L}{R_f + R_L} \tag{2-50}$$

If $R_L >> R_f$,

$$\text{T.U.F.} = \frac{2\sqrt{2}}{2} = 0.287. \tag{2-51}$$

To illustrate what this means, rearrange the T.U.F. equation:

dc power delivered to the load =
ac transformer rating × T.U.F. (2-52)

Consider now a 1 KVA transformer (1000 VA rating). If this were used in a half-wave rectifying circuit, the dc power that could be delivered to a resistive load by this transformer without overheating would be:

1000 × 0.287 = 287 watts.

Actually, due to dc current flowing through the secondary of this transformer and causing saturation effects, the T.U.F. should be reduced even further to 0.2 for a half-wave rectifier.

Fortunately, when filtering circuits are used that cause a more even flow of current the T.U.F. rises close to unity (0.9 for an LC filter with a full-wave rectifier), so that not as much caution is required as in the unfiltered case. But in simple rectifiers, such as battery chargers or welders that usually don't have any filtering, the transformer must be selected with care.

2-7. DISADVANTAGES OF THE HALF-WAVE RECTIFIER

The disadvantages of a half-wave rectifier may be summarized as follows:

(a) Excessive ripple (r = 1.21)
(b) Low ratio of rectification (0.406)
(c) Low transformer utilization factor (0.287)
(d) DC saturation of transformer secondary winding.

2-8. THE FULL-WAVE RECTIFIER

Many of the above disadvantages may be overcome or reduced by the addition of another diode to provide full-wave rectification as shown in Fig. 2-6.

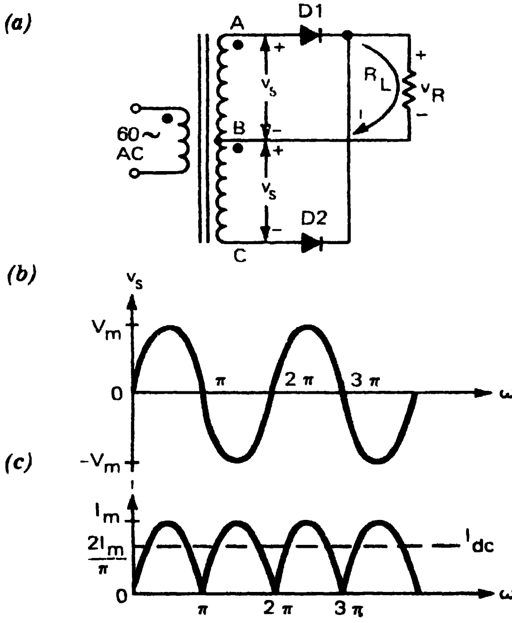


Fig. 2-6. (a) Full-wave rectifying circuit showing center-tapped transformer. (b) Open-circuit voltage from one half of secondary winding. (c) Output current showing the average value, I_{dc}.

In the full-wave rectifier the transformer secondary must be center-tapped and have twice the voltage from line to line (A to C) as the voltage from A to B in the half-wave rectifier. On the positive half cycle of v_s, D1 will conduct, but D2 will be cut off. When

v_s reverses, D2 will take over and "fill in" the gap in the output while D1 is open.

DC Output Current

As for the half-wave rectifier, the peak value of load current is:

$$I_m = \frac{V_m}{R_s + R_f + R_L} \qquad (2\text{-}53)$$

where R_s is now the resistance of one half of the secondary winding.

For an input voltage v_s given by:

$$v_s = V_m \sin \omega t \qquad (2\text{-}54)$$

it can be shown that the output current is represented by:

$$\approx \frac{2I_m}{\pi} - \frac{4I_m}{3\pi} \cos 2\,\omega t - \frac{4I_m}{15\pi} \cos 4\,\omega t. \quad (2\text{-}55)$$

Evidently the average value of output current that a dc ammeter will indicate is:

$$I_{dc} = \frac{2I_m}{\pi} \qquad (2\text{-}56)$$

which is seen to be double that of a half-wave rectifier.

Note that the lowest frequency of the alternating components is now twice the input frequency, so for a 60 Hz input the output ripple frequency is 120 Hz compared with 60 Hz for the half-wave rectifier.

DC Output Voltage

Following a similar discussion to that used for the half-wave rectifier, the dc output voltage across the load is:

$$V_{dc} = \frac{2V_m}{\pi} - I_{dc}\,(R_s + R_f) \qquad (2\text{-}57)$$

and thus at no load the output voltage

$$V_{dc} = \frac{2V_m}{\pi} \qquad (2\text{-}58)$$

or double that of the half-wave rectifier.

Peak Inverse Voltage

Let us determine the maximum reverse voltage that occurs across the diodes in Fig. 2-6(a). Consider v_s on the negative half cycle. D1 will be cut off and D2 conducting. The peak voltage that occurs across R_L is V_m, and at the same time the peak voltage from B to A is also V_m. The sum of these two is the maximum reverse voltage across D1, equal to $2V_m$.

Strictly speaking the term peak inverse voltage, PIV, has been used to denote the maximum reverse voltage *capability* of a diode. But common usage has made this term popular for the maximum reverse voltage that *occurs* across a diode in a rectifying circuit. Consequently, the PIV for a diode in a full-wave rectifying circuit is $2V_m$ compared with V_m in a half-wave rectifier, and is an item to be considered when selecting a diode for a power supply.

Ripple Factor

The ripple factor, r

$$= \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1}\ . \qquad (2\text{-}59)$$

$I_{dc} = \frac{2I_m}{\pi}$, but what is the rms value of i for a full-wave rectifier? The heating effect (which is the definition of rms) of a pulsating full-wave rectified sinusoid is the same as an unrectified alternating sinusoid, so that $I_{rms} = \frac{I_m}{\sqrt{2}}$.

Making these substitutions gives r = 0.482, which is a drastic improvement from 1.21 for the half-wave rectifier.

Ratio of Rectification

Using the previous values for I_{dc} and I_{rms} allows the determination of dc and ac power in the circuit and, not surprisingly, effects a doubling of the rectification ratio from 0.406 for the half-wave rectifier to 0.812 for the full-wave rectifier. (See Problem 2-16.)

Transformer Utilization Factor

In a full-wave rectifier the average T.U.F. is found by considering the primary and secondary windings separately. This yields a value of 0.693 compared with 0.287 for the half-wave rectifier. Also, since current flows in the secondary side in opposite directions, there is no problem due to dc saturation of flux in the core.

2-9. THE BRIDGE RECTIFIER

A commonly used circuit for supplying large amounts of dc power is the bridge rectifier, employing four diodes as illustrated in Fig. 2-7.

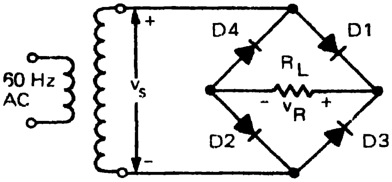


Fig. 2-7. Circuit of the bridge rectifier.

During the positive half cycle of v_s, diodes D1 and D2 conduct, developing the output voltage across R_L. D3 and D4 are reverse-biased. On the negative half cycle D1 and D2 are open, and conduction occurs through D3 and D4, resulting in a full-wave rectified output.

It should be noted that for approximately the same output voltage the transformer secondary line-to-line voltage should be one-half that used for the full-wave rectifier employing two diodes. The dc output current, $I_{dc} = \frac{2I_m}{\pi}$, where

$$I_m \quad R_s + 2R_f + R_L \qquad (2\text{-}60)$$

and the dc output voltage is:

$$V_{dc} = \frac{2V_m}{\pi} - I_{dc}(R_s + 2R_f). \qquad (2\text{-}61)$$

$2R_f$ must be used since two series diodes are conducting at the same time.

The ripple factor and ratio of rectification are the same as for the full-wave rectifier of Section 2-8. However, since the current flowing in the transformer secondary is purely alternating the transformer utilization factor is now increased to 0.812, which is the main reason for the popularity of a bridge rectifier where large amounts of dc power are required.

Another advantage is that the PIV across each diode is the peak voltage across the load = V_m, not $2V_m$ as in the two-diode rectifier. In addition, a center-tapped transformer is not required.

The above advantages easily outweigh the disadvantage of having to use two extra diodes, although this does cause an additional voltage drop that reduces the output voltage and cannot be ignored in low voltage dc power supplies.

Bridge rectifier assemblies are available that contain four diodes encapsulated in a package and have two terminals for connection of the ac input and two for the dc output. Likewise, selenium assemblies are available that have the four rectifiers already interconnected.

2-10. THREE-PHASE FULL-WAVE RECTIFIERS

The bridge rectifier is popular where large amounts of dc power are required because of the higher transformer utilization factor. A circuit which vastly reduces the ripple factor and improves the ratio of rectification is the three-phase rectifier shown in Fig. 2-8.

Consider the interval from $\omega t = \pi/3$ to $\omega t = 2\pi/3$. V_{s1} is positive, which means that current will flow from A through D1 and the load, and back through D5 into B (v_{s2} is negative).

From $2\pi/3$ to π, v_{s3} is at its maximum negative value, which means A is positive with respect to C (See Fig. 2-8(a)). Consequently current will again flow out of A, through D1 and the load, and back through D6 into C. Due to these conduction periods of only 60° the load current never falls below $\sqrt{3}\, I_m/2$.

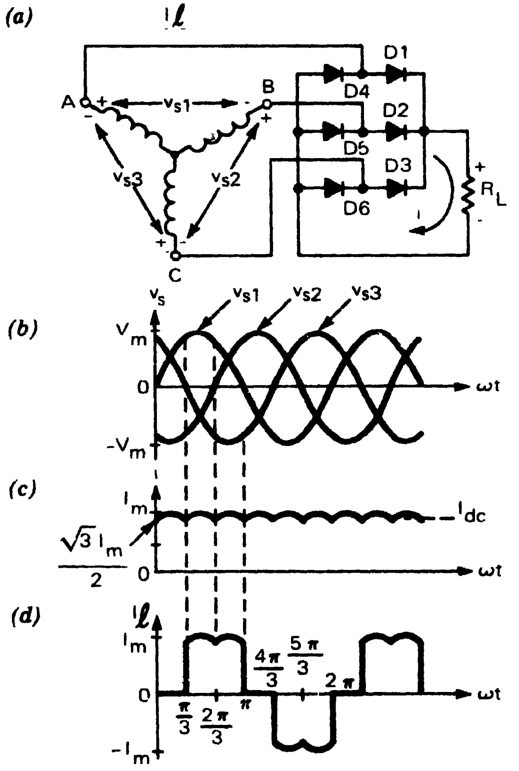


Fig. 2-8. (a) Circuit of a three-phase full-wave rectifier showing the six diodes. (b) Three-phase input voltage, each phase separated by 120°. (c) Load current showing the average value I_{dc}. (d) Line current that flows for each phase (drawn for phase A).

It is evident from Fig. 2-8(c) that the average value of current through the load is higher than any of the previous power supplies and may be shown to be:

$$I_{dc} = \frac{3I_m}{\pi} = 0.955\ I_m \tag{2-62}$$

Similarly, the rms value of the load current is:

$$I_{rms} = 0.957\ I_m \tag{2-63}$$

Using these values, the ripple factor r = 0.055, the ratio of rectification = 0.995, and the transformer utilization factor = 0.955. Small wonder that this is the circuit used to convert ac into dc in the automobile alternator, since there is only 5.5% ripple voltage in the output with no filtering.

In Fig. 2-8(a) the primary windings of the transformer were not shown (in the alternator the windings represent the coils mounted in the stator), but the connection could be made using three single-phase transformers connected in a wye as shown or a delta. In some cases, where even less ripple is required or to reduce the amount of filtering necessary, transformers may be interconnected to give six-phase, nine-phase or even twelve-phase operation, with higher average dc levels of current and voltage resulting. It should be noted that the ripple frequency of the output with three-phase input is now six times the input frequency, a factor that aids in the filtering process. Also, the PIV across each diode is only V_m as in the bridge rectifier.

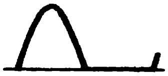
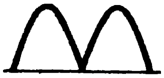
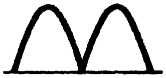
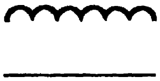
2-11. COMPARISON OF RECTIFIER CIRCUITS

The characteristics of the four rectifiers that we have examined are summarized in Table 2-1.

2-12. READING OF AN AC VOLTMETER

We have considered the readings of a dc ammeter, ac ammeter, dc voltmeter and wattmeter in various places of the rectifying circuits, but not the reading of an ac voltmeter. What an ac voltmeter indicates must be interpreted very carefully, since what the reading means depends upon the type of the voltmeter.

TABLE 2-1 COMPARISON OF RECTIFIER CIRCUITS WITH RESISTIVE LOAD

	HALF-WAVE	FULL-WAVE	BRIDGE	THREE-PHASE FULL-WAVE
Secondary Voltage Line-to-Line ($V_m = \sqrt{2}V$)	V	2V	V	V
Number of Diodes	1	2	4	6
Peak Inverse Voltage	V_m	$2V_m$	V_m	V_m
No Load DC Output, V_{dc}	$\frac{V_m}{\pi} = 0.318\ V_m$	$\frac{2V_m}{\pi} = 0.636\ V_m$	$\frac{2V_m}{\pi} = 0.636\ V_m$	$\frac{3V_m}{\pi} = 0.955\ V_m$
Ripple Factor	1.21	0.482	0.482	0.055
Ratio of Rectification	0.406	0.812	0.812	0.995
Transformer Utilization Factor	0.287	0.693	0.812	0.955
DC Power Available from a 1 KVA Transformer, Watts	287	693	812	955
Ripple Frequency	f	2f	2f	6f
Output Waveform of Voltage and Current				

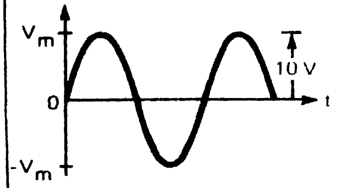
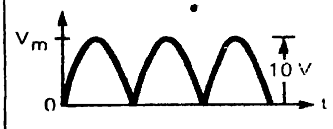
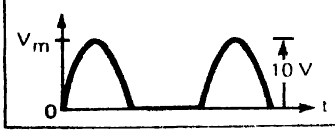
Consider connecting an ac voltmeter across a full-wave rectified output, of peak value V_m. We have seen that the rms value of such a wave is $\frac{V_m}{\sqrt{2}}$, the same as an alternating sinusoid of the same peak value. But, unless the ac voltmeter is a *true rms* meter (which means it responds to the heating value of the wave across which it is connected), the reading will be very different from $\frac{V_m}{\sqrt{2}}$ for the following reason. Many lower priced ac voltmeters convert the ac to a proportional value of dc using a clamping, rectifying and filtering circuit called a voltage doubler. (This is covered in Chapter 4.) Consequently the ac voltmeter responds to the peak-to-peak voltage and is calibrated to read rms values for *true sine waves.* When connected across a full-wave rectified output the voltage doubler circuit inside the voltmeter produces only half the dc voltage that it would for a full sine wave. Thus the voltmeter will indicate $\frac{V_m}{2\sqrt{2}}$, or only half the rms value. It will also indicate this value for a half-wave rectified output since it cannot sense the "gaps" in between the positive peaks.

For some higher priced voltmeters, such as the Hewlett-Packard 410B VTVM, conversion of ac into dc is accomplished using a shunt rectifier with a filter. Such an ac voltmeter will indicate $\frac{V_m}{\sqrt{2}}(1-\frac{1}{\pi})$ on half-wave rectification and $\frac{V_m}{\sqrt{2}}(1-\frac{2}{\pi})$ on full-wave rectification.

These comments are summarized in Table 2-2 where values have been determined for a 10 V peak waveform.

If one attempts to measure the rms ripple voltage in the output of a rectifier, a true rms meter must be used with a capacitor in series with it to block the dc component of voltage, V_{dc}, allowing the meter to respond to only the alternating components.

TABLE 2-2 COMPARISON OF AC VOLTMETER READINGS WITH RMS VALUES

WAVEFORM	TRUE RMS VALUE	HEATHKIT IMW-11 VTVM INDICATION	HEWLETT-PACKARD 410B VTVM INDICATION
	$\frac{V_m}{\sqrt{2}} = 7.07$ V	$\frac{V_m}{\sqrt{2}} = 7.07$ V	$\frac{V_m}{\sqrt{2}} = 7.07$ V
	$\frac{V_m}{\sqrt{2}} = 7.07$ V	$\frac{V_m}{2\sqrt{2}} = 3.53$ V	$\frac{V_m}{\sqrt{2}}(1 - \frac{2}{\pi}) = 2.57$ V
	$\frac{V_m}{2} = 5$ V	$\frac{V_m}{2\sqrt{2}} = 3.53$ V	$\frac{V_m}{\sqrt{2}}(1 - \frac{1}{\pi}) = 4.83$ V

2-13. SUMMARY

(1) The determination of current flowing in a series circuit containing a diode, resistor R_L, and dc applied voltage V is made by drawing a load line on the static characteristic curve of the diode. One end of the load line is at $V_D = V$, $I = 0$, the other at $V_D = 0$, $I = V/R_L$. The intersection of the load line with the curve gives the Q point.

(2) When an alternating voltage is applied to a series diode-resistor circuit, a dynamic curve may be constructed to give the output current waveform for the load resistor for which the curve was drawn.

(3) Ripple factor is defined as:

$$r = \frac{\text{rms value of ac components}}{\text{dc value of the wave}}$$

and is an indication of how successful is a circuit in converting ac to dc.

(4) Ratio of rectification

$$= \frac{\text{dc power delivered to the load}}{\text{ac input power from transformer secondary}}$$

and is a figure of merit (a measure of efficiency) to compare rectifiers.

(5) Transformer Utilization Factor

$$= \frac{\text{dc power delivered to the load}}{\text{ac rating of transformer secondary}}$$

and is used to determine the necessary rating of a transformer for a given dc load.

(6) % Voltage Regulation

$$= \frac{V_{\text{no-load}} - V_{\text{full-load}}}{V_{\text{full-load}}} \times 100\%$$

and is an indication of how the output voltage of a power supply varies with load.

(7) The internal resistance of a power supply,

$R_{INT}\ \ \frac{V_{NL} - V_{FL}}{I_L}$, is due to the resistance of the diodes and transformer winding.

(8) A half-wave rectifier is characterized by an excessive ripple, low values of ratio of rectification and transformer utilization factor and dc saturation of the transformer core. -

(9) The center-tapped and bridge-type full-wave rectifiers improve on the half-wave characteristics but the three-phase, full-wave rectifier is even better, as summarized in Table 2-1.

(10) The reading of an ac voltmeter depends upon the type of meter and the waveform across which it is connected, as summarized in Table 2-2.

QUESTIONS

Q2-1. Explain with the aid of a sketch the use of a diode's characteristic curve to determine the dc current that will flow through a diode when it is connected in series with a resistor and a dc supply.

Q2-2. What does a load line drawn on a diode's characteristic curve represent?

Q2-3. What effectively happens to the load line on a diode's characteristic curve as an alternating voltage is applied?

Q2-4. How does a diode's dynamic characteristic curve differ from its static characteristic curve?

Q2-5. What effect does a diode's cut-in voltage have on the waveform of output voltage compared with the input? How may this effect be removed?

Q2-6. Which power supply has a better voltage regulation, one with 0% or one with 100% voltage regulation?

Q2-7. Is it possible for a power supply to have a voltage regulation in excess of 100%? Explain.

Q2-8. Describe how you would experimentally determine the internal resistance of a power supply.

Q2-9. Describe in your own words what is meant by the statement: "A transformer's utilization factor is 0.693 when used in a full-wave rectifier circuit."

Q2-10. Give four advantages that a full-wave rectifier has over a half-wave rectifier.

Q2-11. What is the maximum reverse voltage, sometimes called the PIV, across a diode in:

(a) a half-wave rectifier
(b) full-wave rectifier
(c) bridge-type full-wave rectifier.

Q2-12. What contributes to the low ripple factor in a three-phase full-wave rectifier?

Q2-13. What would be the ripple frequency in a twelve-phase full-wave rectifier, in terms of the input frequency?

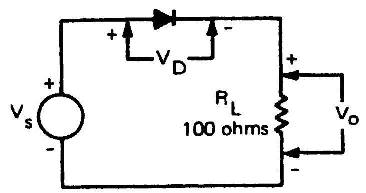
PROBLEMS

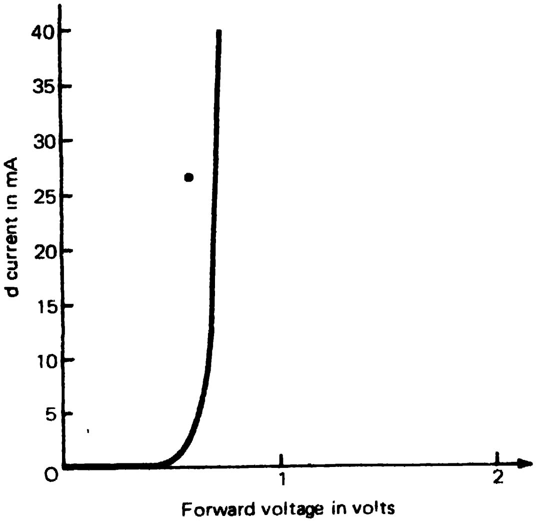
P2-1. For the diode circuit shown on page 31, determine the current flowing when $V_s = 1.5$ V dc is applied. Use the given characteristic of the diode. (Trace onto a clear sheet.)

P2-2. Repeat P2-1 for: (a) $V_s = 2V$, $R_L = 100$ ohms, (b) $V_s = 2V$, $R_L = 50$ ohms. What is the total resistance seen by V_s for the conditions in (a) and (b)? How much of this resistance is due to the diode?

P2-3. Using the circuit and characteristic curve in P2-1, obtain the dynamic characteristic curve of the diode for a 4 V peak-to-peak input and $R_L = 100$ ohms. Then determine the waveform of output current when a 4 V p-p sinusoidal input is applied.

P2-4. Using the dynamic curve from P2-3, repeat P2-3 to determine the current waveforms





for the following conditions:

(a) 1 V p-p sinusoidal input
(b) 4 V p-p square-wave input
(c) 0.5 V p-p square-wave input.

P2-5. (a) Obtain an expression in terms of I_m that approximates the rms value of a half-wave rectified sine wave by using the first two terms of Eq. 2-10 and the fact that the rms value of current is the square root of the sum of the squares of the effective values of each component. The exact value is $\frac{I_m}{2}$.

(b) Repeat using the first three terms.

P2-6. Using $I_m = 1$, *sketch* the first three terms of Eq. 2-10 (a couple of cycles), with proper amplitudes and frequency, and add them algebraically. Is the result close to a half-wave rectified sine wave? Can you see how higher harmonics would give a closer representation?

P2-7. A half-wave rectifier consists of a diode having a dynamic resistance of 1 ohm at its operating current and a transformer whose open-circuit secondary voltage is 12.6 V, 60 Hz; It has a secondary resistance of 3 ohms.

(a) What is the no-load dc output voltage of the rectifier?
(b) What is the dc output voltage when full load draws a dc current of 100 mA?
(c) What is the percentage voltage regulation of this power supply?
(d) What is the internal resistance of this power supply?

P2-8. A power supply has a no-load voltage of 12 V. What will be its full-load voltage if its voltage regulation is:

(a) 10%? (b) 50%? (c) 100%? (d) 200%?

P2-9. If a power supply has a full-load voltage of 20 V, what will be its no-load voltage if its voltage regulation is:

(a) 0%? (b) 25%? (c) 100%?

P2-10. (a) A dc power supply is known to have a ripple factor of 10%. If the dc output voltage is 10 V, what is the rms value of ripple voltage in the output?

(b) Assuming this ripple is sinusoidal in nature, what is its peak-to-peak voltage?

(c) Sketch the waveform of voltage that a dc-coupled oscilloscope would indicate from parts (a) and (b).

P2-11. (a) What is the necessary ac input power from the transformer secondary used in a half-wave rectifier to deliver 500 W of dc power to the load?

(b) What would be the ac input power for the same load in a full-wave rectifier?

P2-12. Given that the total ac input power from the transformer secondary is the sum of the dc power in the load and the ac ripple power in the load, use the fact that the ripple factor in a half-wave rectifier is 1.21 along with Eq. 2-36 to obtain the ratio of rectification for a half-wave rectifier (0.406).

P2-13. Determine the rating of a transformer to deliver 100 watts of dc power to a load under the following conditions:

(a) half-wave rectifier
(b) full-wave rectifier
(c) bridge-type full-wave rectifier
(d) three-phase full-wave rectifier.

P2-14. Repeat P2-5 for a full-wave rectified sine wave using Eq. 2-55.

P2-15. A 120 V ac 60 cycle voltage is applied to the primary of a 5:1 step-up transformer whose secondary is center tapped, allowing a load of 1 K ohm to be connected to a full-wave rectifier utilizing two diodes. Neglecting the voltage drop across the diodes, determine:

(a) The dc voltage across the load.
(b) The dc current through the load.
(c) The dc power delivered to the load.
(d) The VA rating of the transformer secondary.
(e) The ac input power to the transformer, assuming an 80% efficient transformer and a ratio of rectification of this circuit of 0.812.
(f) The ac ripple voltage across the load and its frequency.
(g) The reading of an ac voltmeter (that responds to peak-to-peak value) connected across the load.
(h) The PIV across each diode.
(i) If an additional two diodes were used to connect the 1 K load across a bridge rectifier circuit, utilizing the *full* secondary of the transformer, how much dc power could be delivered using a transformer with the rating found in (d)?
(j) What would be the dc voltage across the load for this connection?
(k) What would be the PIV across each diode in the bridge circuit?

P2-16. Repeat P2-12 to determine the ratio of rectification for a full-wave rectifier (0.812), using the fact that $r = 0.482$ along with Eq. 2-36. Or, obtain the ratio by means of the appropriate values in Eq. 2-39 through 2-46.

P2-17. (a) Determine the rms value of voltage on the secondary of a transformer to provide a no-load dc voltage of 9 V when connected to a full-wave bridge rectifier.
(b) If the secondary winding has a resistance of 3 ohms and each diode has a dynamic forward resistance of 1 ohm, what will be the dc output voltage when a 90-ohm load is connected to the power supply?

P2-18. (a) If an automobile alternator must provide a no-load dc voltage (after rectification) of 14 V, what must be the rms line-to-line voltage at the terminals of the alternator windings?
(b) If the dc load current demanded from the alternator-rectifier is 50 A and each of the diodes has an effective resistance of 0.05 ohm at this value of current, determine what the alternator winding voltage must increase to (by increasing the field current) in order to maintain an output voltage of 12 V dc. (Assume the current flows through one *pair* of diodes at a time.)
(c) What would be the rms ripple voltage in the output for the condition in (b) and what is its frequency if the alternator's frequency is 500 Hz?
(d) What should be the minimum VA rating of the alternator to provide the power demanded in (b)?
(e) What would be the approximate reading of an IMW-11 ac VTVM, of the type indicated in Table 2-2, when connected across the output of the alternator-rectifier for the condition in (b)?

3

FILTER CIRCUITS FOR POWER SUPPLIES

3-0. INTRODUCTION

The 121% ripple in the output of the half-wave rectifier and 48% in the full-wave rectifier is more than can be normally tolerated (except in such applications as battery charging, running dc motors, etc.). The filtering of the unwanted ac component to leave the desired dc output can be done in a number of different ways, according to the load demand. Where low ripple output is desired it is obvious that full-wave rectification should be used. The following discussion considers only full-wave filtering, wherein the frequency of the ripple is 120 Hz for a 60 Hz ac line voltage. This is an advantage where either an inductor (choke) is used to prevent the passage of the ripple current (due to its high inductive reactance to ac but quite low resistance to dc), or a capacitor is used to "short" the ripple to ground but leave the dc to appear at the output. Various combinations of L and C are also used.

3-1. THE INDUCTOR FILTER

Fig. 3-1(a) indicates that an inductor filter simply consists of a choke in series with the load. Since the characteristic of an inductor is to oppose any change in current it will have a smoothing effect as illustrated in Fig. 3-1(b).

In order to determine the ripple and its relation to L, recall that the current flowing in a full-wave rectifier with no filtering may be represented by a series of terms given in Section 2-8. Assuming the third and higher terms contribute little output (especially when the fourth harmonic frequency and higher are effectively blocked due to L), the output current is

$$i \approx \frac{2I_m}{\pi} - \frac{4I_m}{3\pi}\cos 2\omega t. \qquad (3\text{-}1)$$

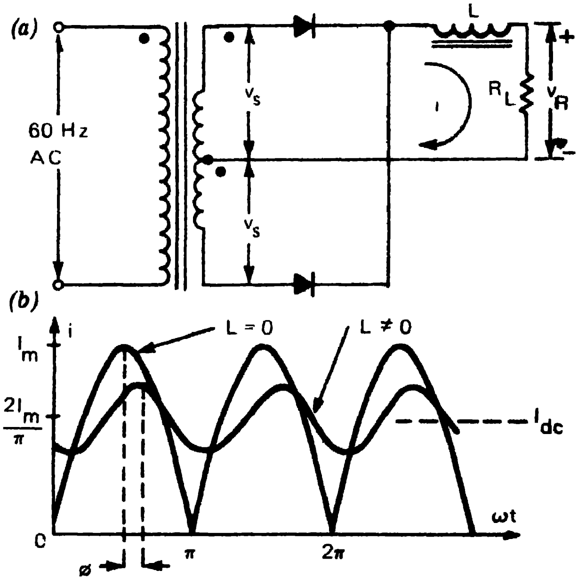


Fig. 3-1. (a) Full wave rectifier with inductor filter. (b) Graph showing load current (and voltage) waveform with and without filtering.

For simplification, let us neglect the diode choke and transformer resistances compared with R_L. Then for the dc component of current, $I_m = \frac{V_m}{R_L}$. (V_m is the no-load peak voltage of v_s.) But the impedance of L and R_L in series with the ac component has a magnitude at 2ω of

$$Z = \sqrt{R_L^2 + (2\omega L)^2}$$

$$= \sqrt{R_L^2 + 4\omega^2 L^2}. \qquad (3\text{-}2)$$

Thus, for the ac component,

$$I_m = \frac{V_m}{\sqrt{R_L{}^2 + 4\omega^2 L^2}}, \tag{3-3}$$

and the resulting load current is given by

$$i \approx \frac{2V_m}{\pi R_L} - \frac{4V_m}{3\pi\sqrt{R_L{}^2 + 4\omega^2 L^2}} \cos(2\omega t - \phi), \tag{3-4}$$

where ϕ is the angle by which the load current lags behind the voltage, and is given by

$$\phi = \tan^{-1}\frac{2\omega L}{R_L}. \tag{3-5}$$

Ripple Factor

Using the definition of ripple factor as the ratio of the rms value of the ac component to the dc value of the wave,

$$r = \frac{I_{r,rms}}{I_{dc}} \tag{3-6}$$

$$= \frac{\dfrac{4V_m}{3\pi\sqrt{2}\sqrt{R_L{}^2 + 4\omega^2 L^2}}}{\dfrac{2V_m}{\pi R_L}} \tag{3-7}$$

$$= \frac{2}{3\sqrt{2}} \times \frac{1}{\sqrt{1 + \dfrac{4\omega^2 L^2}{R_L{}^2}}} \tag{3-8}$$

If $\frac{4\omega^2 L^2}{R_L{}^2} >> 1$, then a simplified expression for r is

$$r = \frac{R_L}{3\sqrt{2}\omega L} \tag{3-9}$$

$$= \frac{R_L}{1600L} \tag{3-10}$$

for a 60-Hz supply, L in henrys, R_L in ohms. This expression clearly shows that reduced ripple will occur for a larger value of L, but it also shows that the ripple will increase as the load is reduced (R_L increased). In fact, as R_L increases to infinity at no load, the use of the equation for a more exact value of r yields $r = \frac{2}{3\sqrt{2}} = 0.471$. This is close to the value of 0.482 obtained in Section 2-8, the difference being attributable to the omission of higher harmonics as mentioned. Clearly the inductor filter should only be used where R_L is consistently small, effective filtering only taking place where the load current is high.

Output Voltage

The output voltage $v_R = iR_L$ has the same waveform as the load current, and the dc output voltage will be the dc component of v_R,

$$V_{dc} = \frac{2V_m}{\pi}. \tag{3-11}$$

However this assumes a perfect inductor with no resistance, and since the inductor filter is only effective with low values of R_L the choke's resistance will not usually be negligible. In this case, the dc output voltage is:

$$V_{dc} = \frac{2V_m}{\pi} \cdot \frac{R_L}{R_L + R}, \tag{3-12}$$

where R is the total of the resistance of the choke, the resistance of half of the transformer secondary winding, and the diode resistance.

This equation may be put in the following form:

$$V_{dc} = \frac{2V_m}{\pi} - I_{dc}R, \tag{3-13}$$

which shows how an inductor with a high resistance can cause poor voltage regulation.

The no-load voltage is evidently $\frac{2V_m}{\pi}$, and the value of R may be obtained experimentally by loading the power supply to its rated value of load current I_{dc} and noting the new dc voltage across the load.

It is to be noted that the load current never drops to zero with a suitable inductor. The magnitude and phase of the induced voltage across the inductor are such as to maintain current flow through the diodes and the transformer secondary windings in a much more even manner than with no filtering. Consequently, the transformer utilization factor is improved with inductive filtering, and the ratio of rectification is almost unity.

3-2. EXPERIMENTAL DETERMINATION OF RIPPLE FACTOR

Assume that the output of a power supply is as shown in the graph of Fig. 3-2.

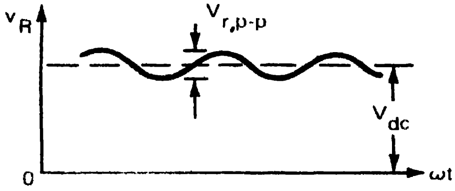


Fig. 3-2. Graph of power supply output voltage showing the ripple voltage and the dc level.

$V_{r,p\text{-}p}$ is the peak-to-peak value of the ripple or ac component, "riding on" a dc level of V_{dc}. This is readily seen on an oscilloscope, or the dc voltage may be read off a dc voltmeter. The ripple voltage will generally not be a sinusoid, but it is converted to the rms value by assuming that it is, except for the capacitor filter as shown later. For example: If $V_{r,p\text{-}p}$ is 10 V and V_{dc} is 300 V, find the ripple factor and the percent ripple:

$$V_{r,max} = \frac{V_{r,p\text{-}p}}{2} = \frac{10}{2} = 5\text{ V}$$

$$V_{r,rms} = \frac{V_{r,max}}{\sqrt{2}} = \frac{5V}{\sqrt{2}} = 3.54\text{ V}$$

$$\text{Ripple Factor } r = \frac{V_{r,rms}}{V_{dc}} = \frac{3.54}{300} = 0.0118$$

$$\%\text{ Ripple} = \text{Ripple Factor} \times 100\%$$
$$= 0.0118 \times 100\% = 1.18\%.$$

3-3. THE CAPACITOR FILTER

We have seen how the inductor filter is only suitable for heavy loads. An inexpensive filter for light loads is found in the capacitor filter which is connected directly across the load, as shown in Fig. 3-3. The capacitor charges up to the peak value of the input voltage and tries to maintain this value, V_m, as the full-wave input drops to zero. The capacitor will discharge through R_L until the input voltage again increases to a value greater than the capacitor voltage. At this point the diode will again be forward biased and there will be a pulse of current through the diode, recharging the capacitor. For light loads (large R_L), the output voltage will remain near V_m. However, as the load increases (R_L decreases), the discharge of C will be greater, resulting in more ripple, and a lower dc output voltage.

3-4. RIPPLE FACTOR, APPROXIMATE METHOD

Referring to Fig. 3-3(d), the average value of load current I_{dc} is the average value of the capacitor discharge current over an interval of T_2. The amount of charge lost by the capacitor during this interval is

$$Q\text{ discharge} = I_{dc} \times T_2\,. \qquad (3\text{-}14)$$

This charge is replaced during a short interval T_1 during which the voltage across the capacitor changes by an amount equal to the peak-to-peak voltage of the ripple, $V_{r,p\text{-}p}$. Using

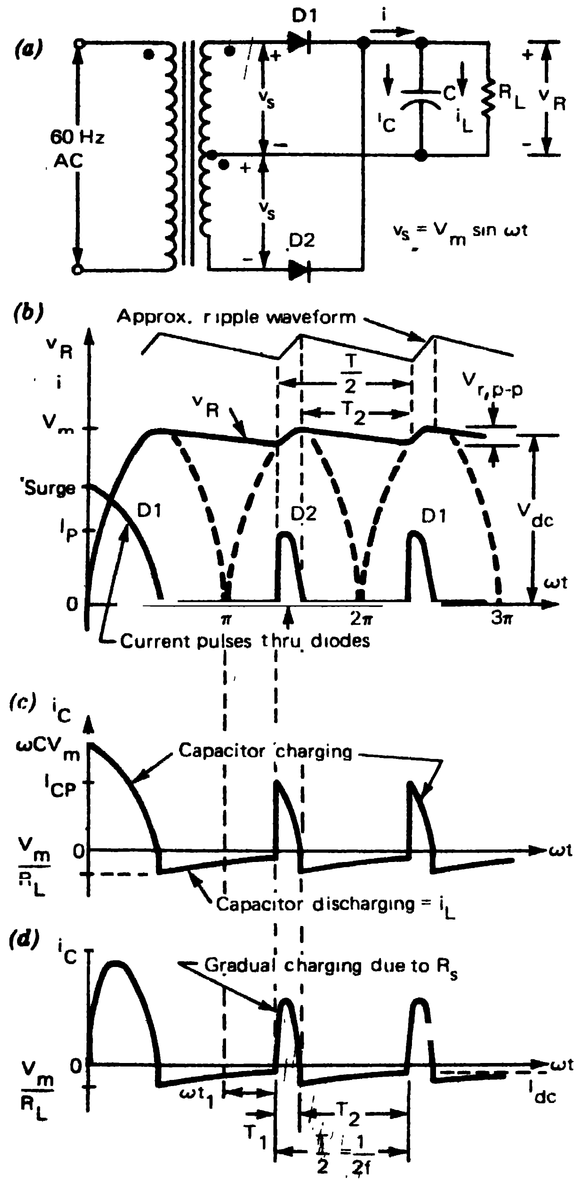

(a)
D1
i
60 Hz AC
v_s
v_s
D2
C
i_C
i_L
R_L
v_R
$v_s = V_m \sin \omega t$
(b)
Approx. ripple waveform
v_R
i
V_m
I_{Surge}
I_P
0
v_R
T/2
T_2
$V_{r,p-p}$
V_{dc}
D1
D2
D1
ωt
π
2π
3π
Current pulses thru diodes
(c)
i_C
ωCV_m
I_{CP}
$\frac{V_m}{R_L}$
0
Capacitor charging
Capacitor discharging = i_L
ωt
(d)
i_C
Gradual charging due to R_s
$\frac{V_m}{R_L}$
0
ωt
I_{dc}
ωt_1
T_1
T_2
$\frac{T}{2} = \frac{1}{2f}$

Fig. 3-3. (a) Full-wave rectifier with capacitor filter. (b) Output voltage waveform and diode current pulses for ideal capacitor. (c) Capacitor charging and discharging current for ideal capacitor. (d) Capacitor current for practical capacitor with significant series resistance, showing the charging and discharging intervals and also the cut-in angle ωt_1.

the relation $Q = VC$, the change in charge is

$$Q \text{ charge} = V_{r,p\text{-}p} \times C. \tag{3-15}$$

But

$$Q \text{ charge} = Q \text{ discharge}, \tag{3-16}$$

$$\therefore V_{r,p\text{-}p} \times C = I_{dc} \times T_2 \tag{3-17}$$

and

$$V_{r,p\text{-}p} = \frac{I_{dc}}{C} \times T_2. \tag{3-18}$$

To this point we have made no approximations. Now let us assume the load is light, the ripple small, and the time for recharging the capacitor small compared with the time for which it discharges. In other words,

$$T_1 << T_2$$

$$\text{and } T_2 \approx \frac{T}{2} = \frac{1}{2f}. \tag{3-19}$$

$$\text{Then } V_{r,p\text{-}p} = \frac{I_{dc}}{2fC}. \tag{3-20}$$

With the assumptions made above, the ripple waveform will be triangular in nature (See Fig. 3-3(b)), with a peak-to-peak value of $V_{r,p\text{-}p}$, and an rms value of

$$V_{r,rms} = \frac{V_{r,p\text{-}p}}{2\sqrt{3}}. \tag{3-21}$$

(This should be remembered when computing the ripple factor experimentally.) Thus, using

$I_{dc} = \frac{V_{dc}}{R_L}$, we obtain:

$$V_{r,rms} = \frac{I_{dc}}{4\sqrt{3}fC} \tag{3-22}$$

$$= \frac{V_{dc}}{4\sqrt{3}fR_LC}, \tag{3-23}$$

and since ripple factor $r = \frac{V_{r,rms}}{V_{dc}}$, we have

$$= \frac{1}{4\sqrt{3}fR_LC} \tag{3-24}$$

$$= \frac{2410}{CR_L} \tag{3-25}$$

for $f = 60$ Hz, C in μF, and R_L in ohms. With these approximations the dc output voltage is

$$V_{dc} = V_m - \frac{V_{r,p\text{-}p}}{2}. \tag{3-26}$$

But $V_{r,p\text{-}p} = \frac{I_{dc}}{2fC}$ and $I_{dc} = \frac{V_{dc}}{R_L}$, from which

$$V_{dc} = V_m - \frac{I_{dc}}{4fC} \tag{3-27}$$

$$= \left(\frac{4fR_LC}{4fR_LC + 1}\right) V_m . \tag{3-28}$$

Thus the ripple may be decreased by increasing C or R_L (or both) with a resulting increase in dc output voltage.

Example 3-1. A laboratory measurement of the output voltage across a load $R_L = 100$ ohms with a filter capacitor $C = 1050$ microfarads is shown in Fig. 3-4. The power line frequency was 60 Hz. Determine the ripple factor and dc output voltage from the graph and compare with calculated values.

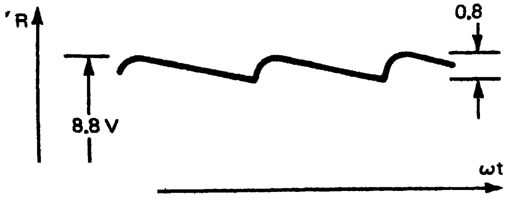


Fig. 3-4. Output voltage waveform observed on oscilloscope for capacitor filter. R_L *= 100 ohms, C = 1050* μ*F.*

Solution. From the graph, $V_{r,p\text{-}p} = 0.8$ V

$$V_{r,rms} = \frac{0.8}{2\sqrt{3}} = 0.231 \text{ V}$$

$$V_{dc} = 8.8 - 0.4 = 8.4 \text{ V}$$

$$r = \frac{V_{r,rms}}{V_{dc}} = \frac{0.2}{8.4} = 0.0238 \text{ or } 2.38\%.$$

Theoretical expected values:

$$r = \frac{2410}{CR_L} = \frac{2410}{1050 \times 100} = 0.023 \text{ or } 2.3\%$$

$$V_{dc} = \frac{4fR_LC}{4fR_LC + 1} V_m$$

$$= \frac{V_m}{1 + \frac{1}{4fR_LC}} = \frac{8.8}{1 + \frac{1}{4 \times 60 \times 100 \times 1050 \times 10^{-6}}}$$

$$= \frac{8.8}{1 + 0.0397} = 8.46 \text{ V}.$$

These are good comparisons, so we can assume the charging time of the capacitor was quite short and made our assumption valid. Note the size of capacitor needed to give this low value of ripple. This was of course due to the quite heavy load of 100 ohms, which drew a current I_{dc} of 84 mA.

Example 3-2. Now consider calculating the ripple factor for an output voltage that was 9 V peak and had a peak-to-peak ripple voltage of 5 V. This was obtained for a load of R_L = 100 ohms, C = 50 microfarads.

Solution.

$$r = \frac{V_{r,rms}}{V_{dc}} = \frac{\frac{V_{r,p\text{-}p}}{2\sqrt{3}}}{V_{dc}} = \frac{\frac{5}{2\sqrt{3}}}{6.5} = 0.222 \text{ or } 22.2\%.$$

$$\text{Theoretical } r = \frac{2410}{CR_L} = \frac{2410}{50 \times 100} = 0.482 \text{ or } 48.2\%$$

Evidently the conditions do not allow us to use the approximate expression for r. This is not surprising since the charging and discharging times of the capacitor are approximately equal to each other, as shown in the next section.

3-5. RIPPLE FACTOR, MORE-ACCURATE METHOD*

In Section 3-4 we found

$$V_{r,p\text{-}p} = \frac{I_{dc}}{C} T_2 \qquad (3\text{-}18)$$

and then assumed $T_2 \approx \frac{1}{2f}$.

$$\text{But } T_2 = \frac{1}{2f} - T_1 \qquad (3\text{-}29)$$

$$\text{and } T_1 \approx \frac{\frac{\pi}{2} - \omega t_1}{2\pi f}. \qquad (3\text{-}30)$$

(See Fig. 3-3(d).) Thus we need to know the cut-in angle, ωt_1, where conduction of the diode (and recharging of the capacitor) begins.

The equation of v_s is: $v_s = V_m \sin \omega t$, (3-31)

and at $\omega t = \omega t_1$, $v_s = V_m - V_{r,p\text{-}p}$. (3-32)

Thus $V_m \sin \omega t_1 = V_m - V_{r,p\text{-}p}$, (3-33)

and, since $V_m = V_{dc} + \frac{V_{r,p\text{-}p}}{2}$, we obtain

$$\sin \omega t_1 = 1 - \frac{V_{r,p\text{-}p}}{V_m} = \frac{1 - \sqrt{3}r}{1 + \sqrt{3}r}, \qquad (3\text{-}34)$$

or, assuming $3r^2 << 1$, $\sin \omega t_1 = 1 - 2\sqrt{3}r$ (3-35)

from which,

$$\omega t_1 = \sin^{-1}(1 - 2\sqrt{3}r) \qquad (3\text{-}36)$$

expressed in radians. Substituting this in $T_1 \approx \frac{\frac{\pi}{2} - \omega t_1}{2\pi f}$ and using the result in $V_{r,p\text{-}p} = \frac{I_{dc}}{C}\left(\frac{1}{2f} - T_1\right)$ yields

$$V_{dc} = \frac{4fR_L C V_m}{4fR_L C + \frac{1}{2} + \frac{\sin^{-1}(1 - 2\sqrt{3}r)}{\pi}} \qquad (3\text{-}37)$$

from which,

$$r = \frac{1}{4\sqrt{3}fR_L C} \times \left[\frac{1}{2} + \frac{\sin^{-1}(1 - 2\sqrt{3}r)}{\pi}\right] \qquad (3\text{-}38)$$

or,

$$r = \frac{2410}{CR_L} \times \left[\frac{1}{2} + \frac{\sin^{-1}(1 - 2\sqrt{3}r)}{\pi}\right], \ f = 60 \text{ Hz}. \qquad (3\text{-}39)$$

The solution of this equation for an explicit expression for r is not easy, but the equation may be used as it stands. Note that the part in brackets is a type of correction factor such that if r is small the correction becomes

$\left[\frac{1}{2} + \frac{\sin^{-1}(1-0)}{\pi}\right] = \left[\frac{1}{2} + \frac{\pi}{2}\frac{1}{\pi}\right] = 1$, and $r = \frac{2410}{CR_L}$ as before. On the other hand, if r is large (about 29%), $r \approx \frac{2410}{CR_L}\left[\frac{1}{2} + 0\right] \approx \frac{2410}{2CR_L}$. (3-40)

Let us use this correction factor to see if we can predict more closely the ripple to be expected in Example 3-2.

$$r = 0.482\left[\frac{1}{2} + \frac{\sin^{-1}(1 - 2\sqrt{3}r)}{\pi}\right] \qquad (3\text{-}41)$$

To solve for r we assume a value for r and see if the expression on the right gives us this value. If we try r = 0.2 (we know it must be less than 0.482) we obtain

---

*This section may be omitted with no loss of continuity.

$$0.482\left[\frac{1}{2}+\frac{1}{\pi}\sin^{-1}(1-2\sqrt{3}\times 0.2)\right]$$

$$=0.482\left[\frac{1}{2}+\frac{1}{\pi}\times(18^\circ)\right]$$

$$=0.482\left[\frac{1}{2}+\frac{0.314}{\pi}\text{ radian}\right]$$

$$=0.482\times 0.6=0.289.$$

Since this does not equal the assumed value of 0.2, we must try another value higher than 0.2. (An increase in r on the right decreases r on the left.) When r = 0.26 is used the result is 0.256, which is close enough. This value of ripple of 26% is much closer to the experimentally observed value of 22% and corresponds to a cut-in angle of only 20° when the more accurate equation for ωt_1 is used:

$$\omega t_1=\sin^{-1}\frac{(1-\sqrt{3}r)}{(1+\sqrt{3}r)}. \qquad (3\text{-}42)$$

This equation must be used since $3r^2$ is now not negligible compared with unity at this value of r. An adjustment for this in determining r gives r = 28%.

The above calculation for r is long and tedious and has been carried out for two reasons. First, it has been used to emphasize that calculations based on approximations will be valid only if the assumptions under which the equations were developed are valid. It turned out in the above discussion that the actual ripple was less than the calculated approximate ripple so the error was on the "safe" side. Secondly, it has been used to show that the cut-in angle, ωt_1, the point at which the diode starts to conduct, is very dependent upon the value of C and will affect the diode's peak current requirements.

3-6. DIODE PEAK CURRENT REQUIREMENTS

Refer to the current pulses through the diodes in Fig. 3-3(b). When voltage was first applied it was assumed v_s was zero and increasing, and the capacitor was uncharged. Consequently, since the voltage across the ideal capacitor (no series resistance) is changing at its maximum rate (maximum slope), the charging current into the capacitor is at its maximum value of

$$I_{Cm}=\frac{V_m}{X_C} \qquad (3\text{-}43)$$

$$=\omega CV_m, \qquad (3\text{-}44)$$

and any time afterwards in the next quarter cycle is given by

$$i_C=I_{Cm}\cos\omega t. \qquad (3\text{-}45)$$

$$=\omega CV_m\cos\omega t. \qquad (3\text{-}46)$$

After the capacitor has completely charged through D1 and subsequently partially discharged through R_L, it is recharged through D2 starting at ωt_1. The peak charging current will occur at this angle and is given by

$$I_{CP}=\omega CV_m\cos\omega t_1 \qquad (3\text{-}47)$$

for an ideal capacitor.

Meanwhile, the diode must also handle the current i_L for R_L, which was previously supplied by the discharging capacitor. The load current is in phase with the sinusoidally increasing input voltage and is given by

$$i_L=\frac{V_m}{R_L}\sin\omega t\text{ for }\omega t_1\leq\omega t\leq\frac{\pi}{2}. \qquad (3\text{-}48)$$

That is, the relation for i_L is only valid for the conduction period of the diode, which is assumed to cut out at the peak of the input voltage. This is not exactly true, but is sufficient for our needs in determining peak diode current.

The total current through the diode is

$$i = i_C + i_L \tag{3-49}$$

$$= \omega C V_m \cos \omega t + \frac{V_m}{R_L} \sin \omega t, \tag{3-50}$$

which applies for the period $\omega t_1 \leq \omega t \leq \frac{\pi}{2}$. It can be seen that i is the sum of a term (i_C) that is decreasing with ωt and a term (i_L) that increases with ωt. Thus, the maximum value of diode current will occur some time after ωt_1 and will not coincide with the peak of the capacitor charging current (although it will be close to ωt_1 when i_L is small compared with i_C). See Fig. 3-3(b).

The angle for the peak diode current is not readily available from the previous equation, because of the limitation on the angle for which this curve is defined. However, we do have an expression for ωt_1 in terms of the ripple r,

$$\sin \omega t_1 = \frac{1 - \sqrt{3r}}{1 + \sqrt{3r}}. \qquad \text{(Equation 3-34)}$$

This will allow us to determine the peak charging current into the capacitor and will give us an idea of the peak diode current. Using $I_{CP}{}^2 = (\omega C V_m)^2 \cos^2 \omega t_1$, $\cos^2 \omega t_1 = 1 - \sin^2 \omega t_1$ and $\sin^2 \omega t_1 = \left(\frac{1 - \sqrt{3r}}{1 + \sqrt{3r}}\right)^2$, we obtain

$$I_{CP} = \omega C V_m \frac{2\sqrt[4]{3}\sqrt{r}}{1 + \sqrt{3r}} \tag{3-51}$$

$$= \omega C V_m \frac{2.63\sqrt{r}}{1 + \sqrt{3r}}, \tag{3-52}$$

at which point the current through the diode for the load resistor will be

$$i_L = \frac{V_m}{R_L} \sin \omega t_1 \tag{3-53}$$

$$= \frac{V_m}{R_L} \frac{1 - \sqrt{3r}}{1 + \sqrt{3r}}. \tag{3-54}$$

The sum of I_{CP} and i_L will give a very close approximation to the peak current, or, to be on the safe side, the peak of the load current, $\frac{V_m}{R_L}$, could be used along with I_{CP}.

Example 3-3. Determine the repetitive peak diode current for the load and filter in Example 3-2 using the previous equations. Given: $R_L = 100$ ohms, $C = 50\ \mu F$, $V_m = 9$ V, $r = 0.28$, $f = 60$ Hz.

Solution.

$$I_{CP} = \omega C V_m \frac{2.63\sqrt{r}}{1 + \sqrt{3r}} = 0.17 \times \frac{2.63\sqrt{0.2}}{1.485}$$

$$= 0.17 \times 0.935 = 0.16 \text{ A}$$

$$i_L = \frac{V_m}{R_L} \frac{1 - \sqrt{3r}}{1 + \sqrt{3r}} = \frac{9}{100} \times \frac{0.515}{1.485} = 0.031 \text{ A}$$

Thus the peak diode current $I_P \approx I_{CP} + i_L = 0.16 + 0.031 = 0.191$ A. It is safe to say that it is no more than $0.16 + 0.09 = 0.25$ A.

The peak voltage measured across a 0.6 ohm resistor in series with the diode of Example 3-3 was 0.13 V, indicating a peak current of 0.216 A, which compares well with our expected values.

Example 3-4. Now determine the repetitive peak diode current for the load and filter of Example 3-1. Given: $R_L = 100$ ohms, $C = 1050\ \mu F$, $V_m = 8.8$ V, $r = 0.023$, $f = 60$ Hz.

Solution

$$I_{CP} = \omega C V_m \frac{2.63\sqrt{r}}{1 + \sqrt{3r}} = 3.48 \times 0.384 = 1.34 \text{ A}$$

$$i_L = \frac{V_m}{R_L} \frac{1 - \sqrt{3r}}{1 + \sqrt{3r}} = \frac{8.8}{100} \times 0.922 = 0.081 \text{ A}$$

$\therefore I_P \approx I_{CP} + i_L = 1.34 + 0.081 = 1.421$ A

This compares with 0.4 A determined experimentally. Why the difference?

3-7. EFFECT OF CAPACITOR SERIES RESISTANCE

We have assumed a perfect capacitor and have not taken into account the series resistance, which can amount to as much as 25 ohms in some electrolytic capacitors, the ones usually used to provide this large amount of capacitance. The reactance of C = 1050 μF at f = 60 Hz is only 2.53 ohms. The current flow is thus determined primarily by the capacitor's series resistance. Assuming this resistance, R_s, is 25 ohms, we effectively have R_s in parallel with R_L or 20 ohms. The peak current then will be $\frac{V_m}{20} = \frac{8.8}{20} = 0.44$ A, very close to that measured.

The series resistance has two other effects. It causes a more gradual charging of the capacitor (See Fig. 3-3(d)), and it also increases the ripple voltage, as shown in Fig. 3-5. The total voltage across the capacitor is made up of $i_C R_s$ and the voltage across the capacitor proper. But when the current i_C starts to decrease after V_m is reached, the voltage to the right of the diode also decreases, allowing the diode to remain forward biased. Not until i_C drops to zero does diode current cease.

To summarize, the diode must be selected for a repetitive peak current, I_P, as well as being able to handle an average current equal to the average load current, I_{dc}. In addition, it is subjected to a surge current when voltage is first applied (since an uncharged capacitor acts as a short circuit initially), that may be equal to $\frac{V_m}{R + R_s}$, where R is the diode and transformer resistance and R_s is the capacitor's series resistance.

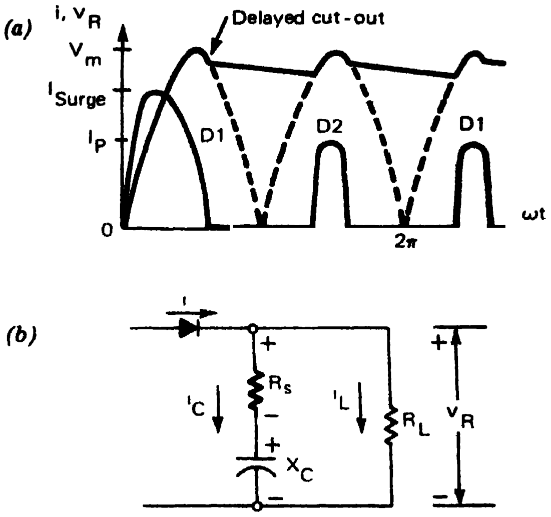


Fig. 3-5. (a) Output voltage waveform showing the delayed cut-out of diode current. (b) Partial diagram of rectifier circuit showing the series resistance of the filter capacitor.

3-8. PIV OF A HALF-WAVE RECTIFIER

Although only full-wave rectifiers are being used in this discussion, it is important to consider the effect of the capacitor filter on the PIV of the diode in a half-wave rectifier. It was shown earlier that the PIV is V_m in a half-wave and $2V_m$ in a full-wave rectifier. But the addition of a capacitor filter to a half-wave rectifier causes the PIV of the diode to be also $2V_m$. This is because the voltage across R_L is held by the capacitor at approximately V_m when the input voltage reverses, instead of falling to zero as is the case with no filter. The reader should draw a diagram to satisfy himself that this is true.

3-9. THE CHOKE INPUT OR LC FILTER

It was found that ripple increased with an increase in load resistance in an inductor filter, but decreased in a capacitor filter for an increase in R_L (C constant). A combination of these two filters into a choke-input or L-section filter should then make the ripple independent of load resistance. An LC filter with a full-wave rectified input is shown in Fig. 3-6, along with the dc output voltage characteristic for varying load current.

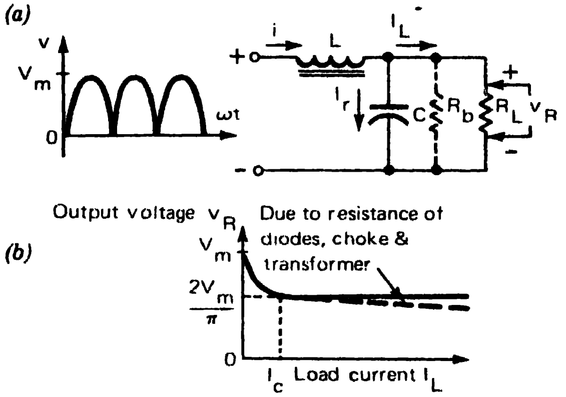


Fig. 3-6. (a) LC filter showing the full-wave rectified input voltage and the bleeder resistor R_b. (b) Graph of dc output voltage as a function of dc load current.

If L is small, the capacitor will be allowed to charge up to V_m and cut off the diodes, allowing only short pulses of current. As L is increased, the pulses of current are smoothed and made to flow for a longer period but at a reduced amplitude until at some critical value of inductance, L_c, either one diode or the other is always conducting, with the result that the input voltage v and current i are full-wave rectified sine waves. This was not the case in the capacitor filter and made it very difficult to analyze, because no mathematical expression representing the input voltage and current was available. Thus, providing we use the critical value of inductance as a minimum (to be determined), the following is an approximation of the voltage applied to the filter and load:

$$v = \frac{2V_m}{\pi} - \frac{4V_m}{3\pi} \cos 2\,\omega t. \tag{3-55}$$

Let us first assume ideal components so that the inductor has no dc resistance. Then the dc output voltage will be $V_{dc} = \frac{2V_m}{\pi}$. If R represents the series resistance of transformer, diodes, and inductor, then the dc output voltage is

$$V_{dc} = \frac{2V_m}{\pi} - I_{dc}R \tag{3-56}$$

and varies as shown in Fig. 3-6(b). This is the same expression as for the inductor filter but the ripple is very different as now shown.

3-10. RIPPLE FACTOR IN AN LC FILTER

The ripple current that does pass through L is not allowed to develop much ripple voltage across R_L (as in a pure inductor filter) if the reactance of C at the ripple frequency is small compared with R_L. Thus, for a properly designed LC filter,

$$X_C << R_L \text{ and } X_L >> X_C \text{ at } 2\omega = 4\pi f. \tag{3-57}$$

Consequently, the ac component of current through L is determined primarily by $X_L = 2\omega L$, and the rms value of ripple current is

$$I_{r,rms} = \frac{4\,V_m}{3\pi\sqrt{2}X_L} \tag{3-58}$$

$$= \frac{2}{3\sqrt{2}X_L} \times \frac{2V_m}{\pi} \tag{3-59}$$

$$= \frac{\sqrt{2}V_{dc}}{3X_L} \tag{3-60}$$

Although we said X_C was small compared with R_L, it is not zero, and the ripple voltage in the output is developed by $I_{r,rms}$ flowing through X_C:

$$V_{r,rms} = I_{r,rms} X_C \tag{3-61}$$

$$= \frac{\sqrt{2}X_C}{3X_L} V_{dc}. \tag{3-62}$$

$$\therefore \text{Ripple Factor } r = \frac{V_{r,rms}}{V_{dc}} \tag{3-63}$$

$$= \frac{\sqrt{2}X_C}{3X_L} \tag{3-64}$$

$$= \frac{\sqrt{2}}{3} \times \frac{1}{2\omega C} \times \frac{1}{2\omega L} \tag{3-65}$$

where $\omega = 2\pi f$, with f = line frequency. At $f =$ 60 Hz:

$$r = \frac{0.83}{LC} \tag{3-66}$$

with L in henrys, C in microfarads, which verifies the independence of r from R_L.

3-11. VALUE OF THE CRITICAL INDUCTANCE

All of our calculations have been based on a critical value of inductance L_c having been selected for L so that the current i does not fall to zero. Since i consists of a dc component of current $I_{dc} = \frac{V_{dc}}{R_L}$ and a sinusoidally varying ac component of peak value $\frac{4V_m}{3\pi X_L}$, we must ensure that I_{dc} exceeds the negative peak of the ac component. That is:

$$\frac{{}_{dc}}{{}_{L}} \geq \frac{4V_m}{3\pi X_L} \tag{3-67}$$

$$\text{or } \frac{V_{dc}}{R_L} \geq \frac{2}{3X_L} V_{dc}, \tag{3-68}$$

$$\text{from which } X_L \geq \frac{2R_L}{3} \tag{3-69}$$

$$\text{or } L_c \geq \frac{R_L}{3\omega}. \tag{3-70}$$

For 60 Hz,

$$L_c \geq \frac{R_L}{1130} \tag{3-71}$$

with R_L in ohms, L in henrys. Due to the approximation made in the series used to represent v, L_c should be made about 25 percent larger:

$$L_c \geq \frac{R_L}{900}. \tag{3-72}$$

It should be noted that when the load is light, for example at no load (R_L large), a large value of L_c is required if the ripple is not to be excessive. However, when R_L decreases (as the load increases), the *critical* value of L_c is also reduced, although this does *not* say that L_c *must* be reduced. This means that the choke inductance should have a large value at no load and may be allowed to decrease as the dc load current increases. Reactors whose inductance does decrease when the dc current through them increases are called "swinging chokes." For example, a choke may have an inductance of 30 henrys at zero current, but only 4 henrys at 100 mA. Of course, it would be nice if the choke could also have an inductance of 30 henrys at 100 mA, but such a constant choke would be bulky and expensive, because it would need more iron to avoid saturation at larger currents.

It can be seen from the curve of output voltage vs. load current in Fig. 3-6(b) that at no load the output is V_m, while at a small current, I_c, the voltage drops to $2V_m$ If L has been made larger than the minimum critical value of inductance, L_c, or if a swinging choke is being used, then a bleeder resistor R_b may be connected across C. (See Fig. 3-6(a).) This will cause at least a minimum current to flow at all times and will prevent the jump in load voltage when the load is removed. For this:

$$R_b \leq 900L_{max}\,. \tag{3-73}$$

The bleeder resistor also provides a path for the capacitor to discharge when the power supply is turned off. The selection of components for an LC filter is given in Example 3-5.

Example 3-5. Design a full-wave rectifier with an LC filter to provide 9 V dc at 100 mA with a maximum ripple of 2%.

Solution.

$$r = \frac{0.83}{LC} \qquad LC = \frac{0.83}{0.02} \approx 42$$

The effective value of $R_L = \frac{V_{dc}}{I_{dc}} = \frac{9}{0.1} =$ 90 ohms

$$L_c \geq \frac{R_L}{900} \quad \frac{90}{900} = 0.1 \text{ henry.}$$

If L is selected to be 0.1 H, then $C = \frac{42}{0.1}$ = 420 μF. However, if L is selected to be 1 H, then C = 42 μF, with a standard value of 50 μF being used in practice.

To maintain good voltage regulation if the load is removed use a bleeder resistor, $R_b \leq 900L_{max}$ 900 × 1 = 900 ohms. Since this is much larger than R_L it will waste little power and is an advantage of using an $L > L_c$.

However, a factor to take into account in choosing the *transformer* is the dc resistance of the choke, which will be larger for a higher value of L.

Let us assume the 1 H choke has a dc resistance of 50 ohms and can handle 100 mA. This will cause a dc voltage drop of 5 V across the choke. Added to this is the drop due to the transformer resistance, which we shall not include here, but it will generally not be negligible. Since $V_{dc,\,(total)} = \frac{2V_m}{\pi}$, then $V_m = \frac{\pi}{2}(9 + 5) = 22$ V. Thus the peak transformer voltage must be 22 V, and a transformer with a voltage from line to center tap of 15.5 V is needed. The PIV of the diodes would be $2V_m$ or 44 V minimum, with an average current capability of 100 mA and a peak current of $I_{dc} + \frac{2}{3X_L} V_{dc} = 100 + 8 = 108$ mA (from Equation 3-60).

This example should be compared with Example 3-4, where to obtain a dc voltage of 8.4 V with a ripple of 2.3% across a load of 100 ohms required a capacitor filter of 1050 μF and a diode peak current of 1.4 A! It should be clear that the smoothing effect of the inductor much improves the transformer utilization factor compared with the capacitor filter. The greater flexibility in selecting filter components and the reduced diode peak current requirements make this type of filter much preferred for load currents in excess of 50 mA where the bulk and expense of the inductor can be tolerated.

3-12. MULTIPLE LC FILTERS

More complete filtering may be obtained by using two or more LC sections as shown in Fig. 3-7.

Analysis proceeds as in the previous section with the resulting ripple factor having a similar form:

$$r = \frac{\sqrt{2}}{3}\,\frac{X_{C1}}{X_{L1}}\,\frac{X_{C2}}{X_{L2}} \tag{3-74}$$

$$= \frac{1.46}{L^2C^2} \tag{3-75}$$

for f = 60 Hz and equal sections (L : L_1 L_2, $C = C_1 = C_2$), with L in henrys, C in μF. Also,

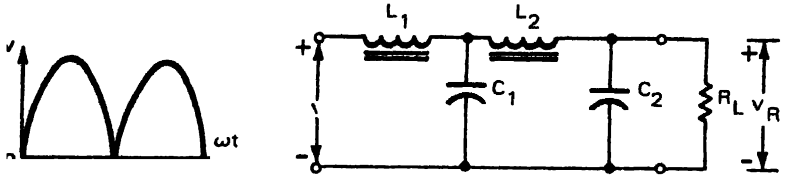


Fig. 3-7. Circuit of a double-section LC filter showing a full-wave input.

$L_c \geq \frac{R_L}{900}$, as for a single section.

3-13. THE CLC OR π FILTER

Where the additional dc voltage drop across L_1 cannot be tolerated in the circuit of Fig. 3-7, a CLC filter will give a similar magnitude of ripple, but with poorer voltage regulation.

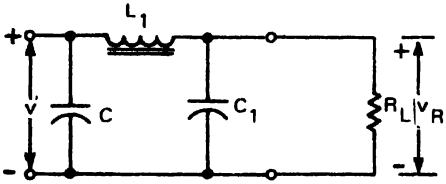


Fig. 3-8. CLC or π-type filter.

Fig. 3-8 shows that the filter basically consists of a capacitor filter followed by an LC section. Consequently v consists of a triangular waveform riding on a dc level, meaning that this filter brings back the problem of high peak diode current pulses, poorer regulation and a ripple dependent upon the load R_L

$$r = \sqrt{2}\,\frac{X_C}{R_L}\,\frac{X_{C1}}{X_{L1}} \qquad (3\text{-}76)$$

$$= \frac{3300}{C C_1 L_1 R_L} \qquad (3\text{-}77)$$

for f = 60 Hz, C in μF, L in henrys, and R_L in ohms. However, the dc output voltage will be closer to V_m, varying under load similarly to the capacitor filter.

It should be noted that under some conditions of light load currents it is possible to replace L_1, with resistor R equal in value to the reactance of L_1 at the ripple frequency of 120 Hz without changing the ripple factor. However, the increased resistance to dc current flow will make this impractical except where the load current is very small.

3-14. COMPARISON OF FILTER CIRCUITS

A summary of the filter characteristics for a full-wave rectifier is shown in Table 3-1. Note that diode, transformer and choke resistance have been assumed negligible in stating the output dc voltage at a load of I_{dc}.

3-15. SUMMARY

(1) An inductor filter has a ripple that increases with load resistance and consequently is used only with relatively high load currents.

(2) The ripple factor may be determined experimentally by observing the output voltage on an oscilloscope and determining the peak-to-peak value of the ripple voltage. This is converted to an rms value assuming it to be a sinusoid except in the case of a capacitor filter, in which the ripple will be more triangular and has an rms value of $\frac{V_{r,p\text{-}p}}{2\sqrt{3}}$. The ratio of this rms value to the dc average of the wave is the ripple factor.

TABLE 3-1 SUMMARY OF FILTER INFORMATION (FULL-WAVE RECTIFIER)

	TYPE OF FILTER				
	NONE	L	C	L-SECTION	π-SECTION
V_{dc}, no load	$0.636V_m$	$0.636V_m$	V_m	V_m (No R_b)	V_m
V_{dc}, load I_{dc}	$0.636V_m$	$0.636V_m$	$V_m - \frac{4{,}170 I_{dc}}{C}$	$0.636V_m$	$V_m - \frac{4{,}170 I_{dc}}{C}$
Ripple factor r	0.48	$\frac{R_L}{1600L}$	$\frac{2{,}410}{CR_L}$	$\frac{0.83}{LC}$	$\frac{3{,}330}{C_1C_2LR_L}$
Peak inverse	$2V_m$	$2V_m$	$2V_m$	$2V_m$	$2V_m$

C is in microfarads, L in henrys, R_L in ohms, V_m in volts, and I_{dc} in amperes. (From *Electronic Devices and Circuits* by Millman and Halkias. Copyright © 1967 by McGraw-Hill Book Company.)

(3) The ripple from a capacitor filter decreases as the load resistance increases, and provides effective filtering only for light loads of 50 mA and less.

(4) The capacitor filter provides poor voltage regulation because of the increase in ripple with load current causing a decrease in average voltage.

(5) When the ripple from a capacitor filter is known to be high, in the region of 30%, a more accurate determination of ripple may be be made using $r = \frac{2410}{2CR_r}$ instead of $r = \frac{2410}{CR_L}$.

(6) A capacitor filter is characterized by the high peak current that the diodes must handle due to the capacitor recharging in a very short interval, and this current increases with the value of C.

(7) A diode's PIV in a half-wave rectifier employing a capacitor filter is $2V_m$.

(8) A choke input or LC filter's ripple is independent of load provided a critical inductance, $L_c \geq \frac{R_L}{900}$ is used at all times.

(9) The use of a bleeder resistor or swinging choke will help maintain good voltage regulation in an LC filter.

(10) Use of an LC filter makes the selection of filtering components very flexible and requires less capacitance than in a capacitor filter for a given amount of ripple. It also reduces the peak diode current requirement and is generally preferred for load currents in excess of 50 mA.

(11) Multiple LC sections or the use of a CLC filter improve the filtering process, and in some cases the inductor may be replaced by a resistor for light loads.

QUESTIONS

Q3-1. What characteristic of a choke is relied upon to provide filtering?

Q3-2. Under what conditions does a choke provide effective filtering?

Q3-3. Explain how you would experimentally determine the ripple factor of a power supply using a dc-coupled oscilloscope.

Q3-4. Explain in your own words the filtering action of a simple capacitor filter, and how a larger capacitor, although giving a smoother output, requires higher-peak-current-capability diodes.

Q3-5. What effect does the series resistance of a capacitor, used in a capacitor filter, have on the diode current and ripple voltage compared with an ideal capacitor of no resistance?

Q3-6. How does ripple voltage vary with increase in load resistance in the following filters?

(a) inductor,
(b) capacitor,
(c) LC or choke input.

Q3-7. An LC filter has a full-wave rectified input of peak value V_m applied to it. If no bleeder resistor is employed, what is the no-load dc voltage?

Q3-8. What two advantages does an LC filter have over a capacitor filter in providing a given dc voltage at a required ripple? What disadvantage does it have?

Q3-9. Does ripple depend upon load resistance in a multiple-LC-section filter?

Q3-10. Under what conditions can the inductor in a π-type CLC filter be replaced by a resistor?

PROBLEMS

P3-1. What should be the value of inductance to use in an inductor filter connected to a full-wave rectifier operating at 60 Hz if the ripple is not to exceed 5% for a 100-ohm load? If you used the simplified expression for r from Eq. 3-7 justify its use.

P3-2. Repeat P3-1 using the standard aircraft power frequency of 400 Hz.

P3-3. If the resistance of the choke in P3-1 is 20 ohms, what will be the dc output voltage with the 100 ohm load, if the *line-to-line* voltage of the transformer secondary is 120 V? (Ignore the resistance of the diodes and transformer.)

P3-4. What is the voltage regulation of the rectifier-filter combination in P3-3?

P3-5. The output of a dc power supply is observed on an oscilloscope and found to have an approximately sinusoidal ripple of 5 V p-p "riding" on a dc level of 15 V.

(a) Sketch this voltage waveform.
(b) Determine the percent ripple.

P3-6. Repeat P3-5 with the ripple triangular in nature.

P3-7. A 12.6 V center-tapped transformer is used in a full-wave rectifier with a 100 μF capacitor to provide a filtering for a 1K ohm resistive load. Determine:

(a) the percent ripple in the output.
(b) the dc output voltage.

P3-8. (a) If the lowest dc voltage that could be tolerated in P3-7 was 6 V, what minimum size of capacitor could be used for the same load?

(b) What would be the percent ripple?

P3-9. (a) What value of load resistance would drop the dc voltage in P3-7 to 6 V with the same capacitor (100 μF)?
(b) What would be the percent ripple?

P3-10. (a) Determine the theoretical ripple factor in a full-wave rectifier operating at 60 Hz with a 100 μF filter capacitor and 100 ohm load, using Eq. 3-20.
(b) Recalculate the ripple using Eq. 3-32.
(c) Which value do you think is more accurate? Justify your answer by calculating the cut-in angle ωt_1 using Eq. 3-35 and comparing with $\pi/2$.

P3-11. A capacitor filter connected to a full-wave rectifier operating at 60 Hz provides a dc output voltage that has a peak value of 24 V and a 20% ripple factor. If the load is 100 ohms and the capacitor 100 μF, determine the repetitive peak diode current assuming an ideal capacitor (negligible series resistance).

P3-12. Repeat P3-11 with

(a) $R_L = 100$ ohms, $C = 500\,\mu F$, $r = 0.05$,
(b) $R_L = 10$ ohms, $C = 5{,}000\,\mu F$, $r = 0.05$.

P3-13. Repeat P3-11 with $R_L = 50$ ohms, $C = 100\,\mu F$. (Note: It will be necessary to first determine r.)

P3-14. A capacitor filter connected to a full-wave rectifier operating at 400 Hz provides a dc output voltage that has a peak value of 12 V. If the load is 500 ohms and the capacitor is 25 μF, determine:

(a) the repetitive peak diode current assuming an ideal capacitor with zero series resistance.
(b) the initial surge current assuming a capacitor with 10 ohms series resistance, and a diode and transformer resistance of 5 ohms.

P3-15. Repeat P3-14 if the load is increased to 50 ohms.

P3-16. (a) What will be the PIV across the diodes in P3-11 and P3-14?
(b) If one of the diodes burned out, what would be the PIV?

P3-17. An LC filter is to be used to provide a dc output with 1% ripple when operating from a full-wave rectifier operating at 60 Hz. To conserve the size of choke, L/C = 0.01 is recommended (L in henrys, C in microfarads). Determine the required values of L and C.

P3-18. Determine the minimum inductance permitted in an LC filter where the minimum load expected is 1 K ohm, and the dc voltage is not to exceed $\frac{2V_m}{\pi}$.

P3-19. Design a full-wave rectifier with an LC filter to provide 12 V dc at 250 mA with a maximum ripple of 0.5%. Specify:

(a) Your suggested *practical* values of L and C.
(b) A bleeder resistor to maintain good voltage regulation.
(c) The transformer secondary voltage, assuming the choke has a dc resistance of 10 ohms.
(d) The PIV capability of the diodes.
(e) The peak current capability of the diodes.

P3-20. It is required to design an LC filter to be connected to the output of a full-wave rectifier which operates at 60 Hz and has an internal resistance of 1 ohm and an open circuit (no-load) voltage (V_m) of 20 V *peak.*

(a) If the output of the filter should have a ripple factor of not more than 0.1%, determine the minimum and maximum inductance values of a swinging choke and the corresponding value of capacitance, if the maximum load is 100 ohms and the minimum load is 1000 ohms.
(b) What should be the maximum dc resistance of the choke if the voltage regulation is not to exceed 2% over the load variation 1000 ohms to 100 ohms?
(c) What would be the dc output voltage under the two extremes of load with this filter connected?
(d) If the choke were accidentally shorted out what would be the peak value of diode current under heaviest load conditions? (Assume an ideal capacitor.)
(e) If both the choke and the output were shorted, what would be the peak diode current? (Neglect resistance of diodes.)

P3-21. If an identical LC section is added to the one determined in P3-17 what would be the new ripple factor?

P3-22. (a) A π-type CLC filter is to be used at 60 Hz to provide 6 V dc output with 0.1%

ripple for a load of 10 K ohms. If the two capacitors are both 50 μF, what must be the minimum value of inductance to use?

(b) What value of resistance could be used to replace the inductance and still provide the same ripple?

(c) What would be the approximate dc output voltage with the resistor as in (b)?

P3-23. Repeat P3-22 for a 1 K ohm load.

4

THE DIODE CLAMPER AND VOLTAGE DOUBLER

4-0. INTRODUCTION

Another source of dc voltage for light loads is the voltage doubler, again both half-wave and full-wave versions exist, but it is possible by adding sections to have a voltage tripler or quadrupler, etc. The basic advantage of these circuits over conventional power supplies is that no transformer is required.

The voltage doubler evolved from a more basic circuit, the clamper. The clamper, together with a filter, is used in many vacuum tube voltmeters to convert ac to dc. With a slight modification the clamper (or shunt rectifier) can be converted to a half-wave voltage doubler, which is also used in VTVMs to convert ac to dc.

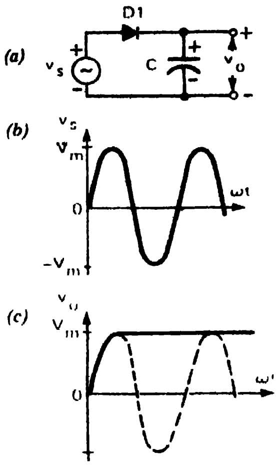


*Fig. 4-1. (a) Half-wave rectifier with a **capacitor as** the load. (b) Waveform of input voltage. **(c) Output** voltage, v_O, across C.*

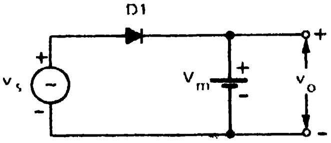


Fig. 4-2. Equivalent circuit of Fig. 4-1(a) with no load connected.

4-1. THE DIODE CLAMPER

Consider a standard half-wave rectifier with a capacitor as the load, as shown in Fig. 4-1(a). With no resistive load across the capacitor, C will charge up to the peak of the input, V_m, as shown in Fig. 4-1(c), and will remain there, acting essentially as a battery source of V_m. (Theoretically, if no discharge of C occurs, D1 will never be forward biased again and no more current will flow in the circuit.) See Fig. 4-2.

Now reverse the positions of the diode and capacitor, as depicted in Fig. 4-3(a). The capacitor will still charge up to V_m, as before, with polarity as shown, and will oppose the input voltage v_s (notice polarity signs are opposite). See Fig. 4-3(a). The output voltage v_o will then be equal, by Kirchoff's Voltage Law to $v_s - v_C$. $\therefore v_o = v_s - V_m$. To obtain $v_s - V_m$ simply shift the curve of v_s "downward" by the amount V_m so that the peak voltage is now *clamped* at zero volts, as shown in Fig. 4-3(c). (This is the same as subtracting a constant value V_m from every point on the sine curve of v_s over a full cycle.)

Notice that v_o is still pulsating but always in a negative direction, so that we have

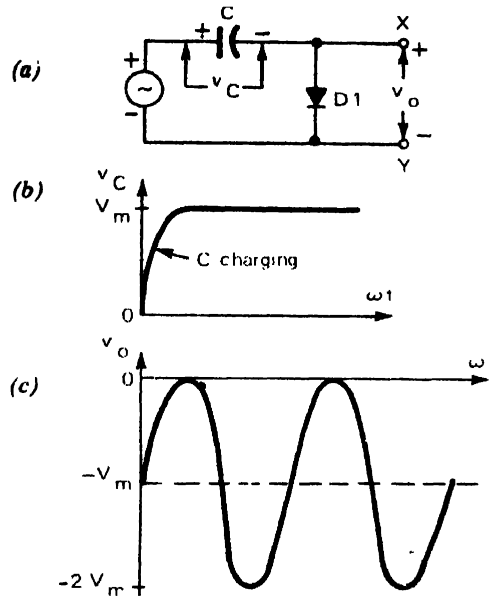


Fig. 4-3. (a) Diode clamper circuit. (b) Capacitor voltage waveform. (c) Output voltage waveform.

effectively rectified the input. (This is why a clamper is often referred to as a shunt rectifier, since it provides a unidirectional output.) Since a dc voltmeter reads the *average value of a wave* there will exist at X a negative voltage with respect to Y and equal in value to V_m. It should be obvious after reversing the direction of the diode (upwards) that the sine wave will again be clamped at zero volts but v_o will now fluctuate only positively, between 0 and $2V_m$, with an average value of $+V_m$.

Example 4-1. Given the input waveform shown in Fig. 4-4(a), determine the outputs from the clamping circuits in Fig. 4-4(b), (d).

Solution. Considering the circuit in Fig. 4-4(b), an ideal diode D1 will permit the capacitor C to charge up to 2 V when v_s goes negative. Thus, $v_o = v_s + V_2 = v_s + 2$, so that the output voltage is the input waveform shifted up 2 volts. This clamps the waveform at 0 V, with a peak value of 12 V. This is shown in Fig. 4-4(c).

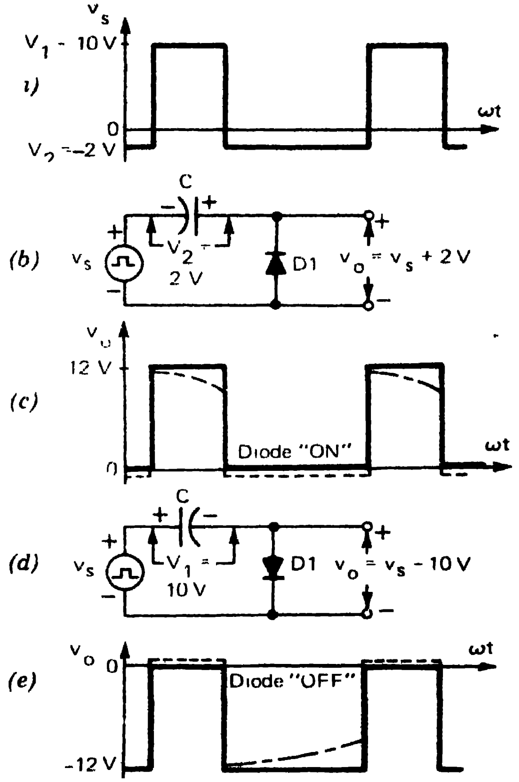


Fig. 4-4.
(a) Input waveform to clamping circuits.
(b) Positive clamping circuit.
(c) Output waveform from circuit of Fig. 4-4(b).
(d) Negative clamping circuit.
(e) Output waveform from circuit of Fig. 4-4(d).

In Fig. 4-4(d), the capacitor charges up to 10 V when v_s is positive and shifts the input waveform down by this amount, as shown in Fig. 4-4(e). That is, $v_o = v_s - V_1 = v_s - 10$.

It should be noted that the voltage drop that occurs across the diode will not give exactly the results shown in Fig. 4-4(c), (e), since the capacitor will not charge up completely to either 2 V or 10 V, but to approximately 1.5 V and 9.5 V. In this case the output will be as shown dotted. For larger voltages the difference can be neglected.

The addition of a load across the diode will allow discharging and will modify these waveforms further, causing "rounding" at the trailing edges, shown by the chain-dashed curves in Figs. 4-4(c), (e).

4-2. SHUNT RECTIFIER WITH FILTER

Now consider adding an RC filter to the output of a shunt rectifier as shown in Fig. 4-5.

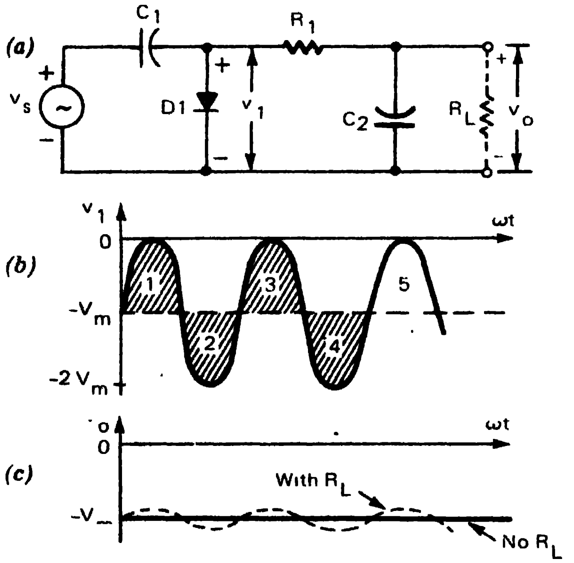


Fig. 4-5. (a) Circuit of a shunt rectifier with an RC filter. (b) Voltage waveform across the diode. (c) Output voltage waveform.

The filter will smooth the shunt rectifier's output, which with a light load will be relatively free from ripple, with a dc value of V_m. This is because the energy stored in the capacitor C_2 during intervals 2 and 4 will "make up" for the "holes" during intervals 3 and 5 to give an average value of V_m. Note that C_2 never charges up to $2\ V_m$ if R_1 has been selected properly. Note also that, if R_1 is zero, the voltage across C_2 will follow the voltage across D1, and no filtering will take place. Although filtering increases with increasing values of R_1, a limit is placed upon R_1 to avoid excessive voltage drop across it, which would give a low output voltage v_o at the load.

4-3. SHUNT RECTIFIER AND FILTER APPLICATION IN A VTVM

This type of circuit is used in the Hewlett-Packard 410B VTVM as shown in Fig. 4-6 with actual values. This circuit is in the ac probe of the meter and is used to convert the ac being measured to a proportional dc voltage. In this way the standard permanent-magnet moving-coil (PMMC) type of movement, designed for dc, may be used to indicate ac values on a properly calibrated scale. A photograph of the H-P 410B VTVM and its ac probe is shown in Fig. 4-7.

The standard high-frequency probe for the Hewlett-Packard 410B VTVM is capable of ac voltage measurements over a range of frequencies from 20 Hz to 700 MHz ±1 dB. This high-frequency capability comes from using a high-frequency diode (vacuum tube) that has a capacitance of only 1.3 pF from plate to cathode and hence removes any shunting effect of the signal to ground over this range. This would be practically impossible to achieve with a semiconductor diode whose diffusion capacitance may be in the order of hundreds of picofarads when forward-biased.

It should be noted that filtering is accomplished by the 22 M ohms resistor and by C_2 and C_3 in parallel, yielding 0.0052 μF. The small capacitor C_2 of 0.0005 μF provides sufficient filtering at high frequencies so that the inductance of the flexible cable is of no consequence and will not affect readings. At lower frequencies additional filtering is necessary and

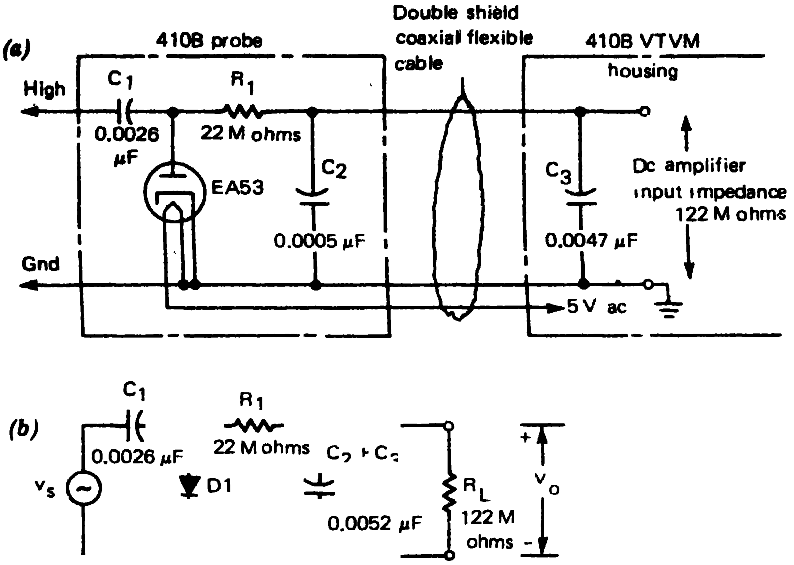

(a)
410B probe
Double shield coaxial flexible cable
410B VTVM housing
High
C1
R1
0.0026 μF
22 M ohms
EA53
C2
0.0005 μF
Gnd
C3
0.0047 μF
Dc amplifier input impedance 122 M ohms
5 V ac
(b)
C1
R1
0.0026 μF
22 M ohms
C2 + C3
0.0052 μF
vs
D1
RL
122 M ohms
+
vo
-

Fig. 4-6. (a) Ac input probe circuit for H-P 410B VTVM. (Courtesy of Hewlett-Packard Co.) (b) Equivalent circuit of probe showing the shunt rectifier and filter.

Fig. 4-7. Photograph of the H-P 410B VTVM and its ac probe. (Courtesy of Hewlett-Packard Co.)

is provided by C_3 of 0.0047 μF in the VTVM housing. The reason for incurring the expense of building the shunt rectifier into the probe is to allow the physical placement of the rectifier as close as possible to the high-frequency voltage being measured to avoid losses due to capacitive and inductive effects in the cable. The dc output of the shunt rectifier and filter will not be affected as it travels up the cable to the VTVM housing. Hence the frequency response of the meter is practically flat from 20 Hz to 700 MHz. However, the dc output voltage will be proportional to the magnitude of the ac input voltage. This dc voltage is applied to a balanced bridge circuit with a meter connected across it, so that a deflection proportional to the ac input is obtained. This is usually *calibrated* to read the true rms value for a pure sine wave input. It should be noted that an ac voltmeter employing a shunt rectifier reacts to the peak of the alternating input waveform. Thus the reading represents true rms values only for pure sine waves, as can be seen by considering the three waveforms in Fig. 4-8.

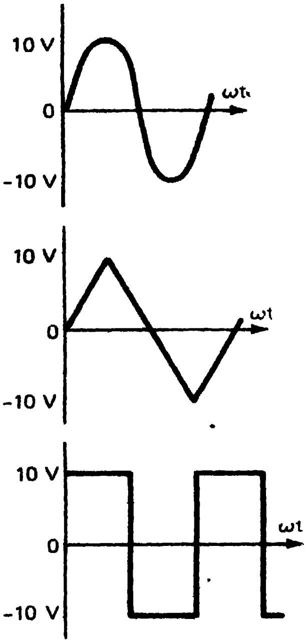


Fig. 4-8. Showing sinusoidal, triangular and square-wave voltage waveforms that have true rms values of 7.07 V, 5.77 V, and 10 V respectively.

For an alternating wave, the VTVM will react to the peak positive voltage and will read the same for all inputs. Since the rms value of the sine wave is 7.07 V the meter will indicate this. However, the rms value for the triangular wave is not 7.07 V (it is 5.77 V), but the meter will indicate 7.07 V, which is not the true value. With the square-wave input the rms value is 10 V, but still the meter will read 7.07 V.

Note that the input impedance to the dc bridge measuring circuit in the 410B is 122 M ohms so that there is negligible loading effect and a high value of resistance for R_1 (22 M ohms) can be used for good filtering. The ac input impedance of the VTVM at lower frequencies is 10 M ohms (shunted by 1.5 pF), and the accuracy is ±3% of full scale on all ranges. The dc input impedance is 122 M ohms on all ranges, and the accuracy is the same ±3%.

4-4. HALF-WAVE OR CASCADE VOLTAGE DOUBLER

If the filter resistor R_1 in the shunt rectifier is replaced by a diode, a voltage doubling action occurs where the output is 2 V_m, as shown in Fig. 4-9. Since the clamped output v_1 has a peak value of 2 V_m, D2 will be forward-biased and allow C_2 to charge to 2 V_m. However, as v_1 starts to decrease again toward zero, D2 will be reverse-biased, and the voltage across C_2 will remain at 2 V_m except for the discharge that will occur through R_L. The ripple frequency will only be 60 Hz due to the half-wave rectification. If semiconductor diodes are used

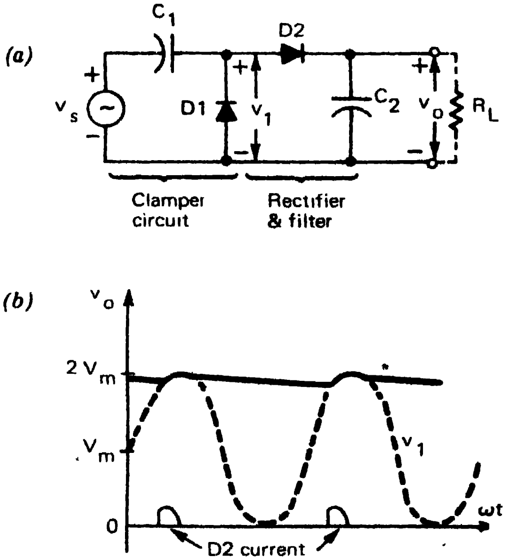


Fig. 4-9. (a) Circuit of a half-wave voltage doubler. (b) Waveform of output voltage.

and if the circuit is used as a power supply with no input transformer, a current-limiting resistor may be placed in series with C_1 to protect the diodes from the surge of charging current when power is first applied.

The half-wave voltage doubler is an inexpensive method of obtaining the high voltage necessary for accelerating purposes in the cathode-ray tube (CRT) of an oscilloscope. A practical circuit to provide a nominal 3 kV ($2\sqrt{2}$ V) at the CRT is shown in Fig. 4-10. Diodes D1 and D2 are selenium, rated at 4 kV and 0.6 mA. The 680 K ohms resistor and 0.01 μF capacitor provide sufficient filtering for the high resistance load presented at the CRT.

4-5. HALF-WAVE VOLTAGE DOUBLER APPLICATION IN A VTVM

The circuit of Fig. 4-10 also finds applications in many lower priced VTVMs where the ac voltage to be measured is converted to a proportional dc voltage. Since this voltage is used to upset the balance in a bridge circuit involving vacuum tubes, it is required that the voltage be negative, so that the diodes in Fig. 4-10 will be reversed to yield $-2\,V_m$. This is shown in the partial diagram of a Heathkit VTVM, Model IMW-11. See Fig. 4-11.

The 6AL5 duodiode is used with two 0.02 μF capacitors as the voltage doubler, which, with further filtering, is shown between X and Y in Fig. 4-11. The resulting negative dc voltage proportional to the ac input voltage is applied (through a range-selector switch) to the grid of a triode (Y). With no input voltage, both grids of the twin-triode 12AU7 are at ground level, and equal current flow from the plates through the cathode resistors places equal potentials on each side of the meter, hence no deflection. (Note: Any variation in plate supply voltage affects both sections the same and does not change the balance.)

Should a negative voltage be applied at Y (whether directly from a dc voltage being measured, or a dc voltage proportional to ac, or a dc voltage proportional to a resistor under measurement) the current through the "left-hand" side of the tube will decrease, upsetting the balance (obtained by the zero adjust). The "right-hand" side of the meter will be at a higher potential, and the meter will deflect upscale.

It should be noted that if resistance is being measured the selector switch will automatically reverse the connection to the meter. This is necessary since when the resistance under measurement is infinite (open circuit) there will be no voltage drop across the resistors in the voltage divider in the VTVM, and the full +1.5 V from the battery will be applied to the grid at Y. This will increase the current flow and make the left-hand side of the meter at a higher potential than the right. In this condition the ohms adjust control would be used to make the deflection full scale and this point would be calibrated as infinity. Any resistance connected at the probe would cause a voltage drop in the series divider, and the voltage at

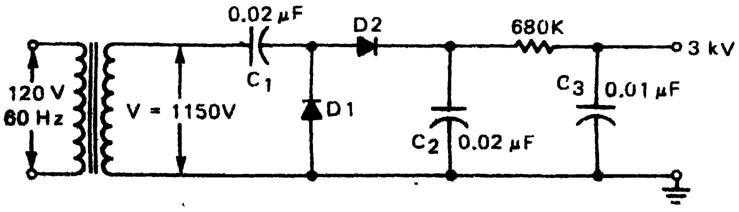
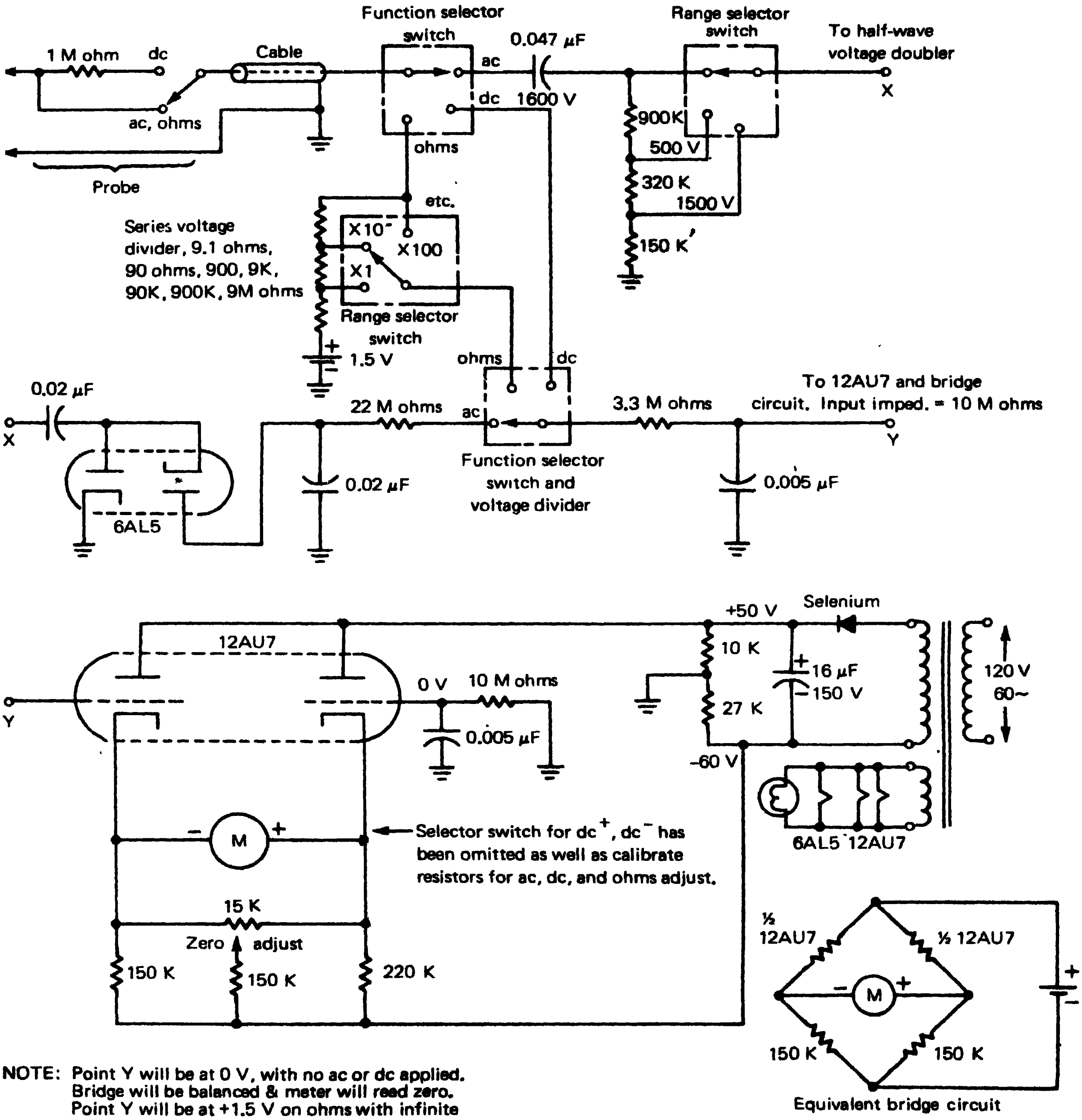


Fig. 4-10. Half-wave voltage doubler to provide 3 kV at CRT.


1 M ohm
dc
Cable
ac, ohms
Probe
Function selector switch
ac
dc
ohms
0.047 µF
1600 V
Range selector switch
To half-wave voltage doubler
X
900K
500 V
320 K
1500 V
150 K
etc.
X10
X100
X1
Series voltage divider, 9.1 ohms, 90 ohms, 900, 9K, 90K, 900K, 9M ohms
Range selector switch
1.5 V
ohms
dc
0.02 µF
X
6AL5
0.02 µF
22 M ohms
ac
Function selector switch and voltage divider
3.3 M ohms
To 12AU7 and bridge circuit. Input imped. = 10 M ohms
Y
0.005 µF
Y
12AU7
0 V
10 M ohms
0.005 µF
+50 V
Selenium
10 K
27 K
16 µF
150 V
120 V
60~
-60 V
6AL5 12AU7
M
Selector switch for dc⁺, dc⁻ has been omitted as well as calibrate resistors for ac, dc, and ohms adjust.
15 K
Zero adjust
150 K
150 K
220 K
½ 12AU7
½ 12AU7
M
150 K
150 K
Equivalent bridge circuit

NOTE: Point Y will be at 0 V, with no ac or dc applied. Bridge will be balanced & meter will read zero. Point Y will be at +1.5 V on ohms with infinite resistance and meter will deflect to full scale since bridge is unbalanced.

the grid would drop from +1.5 V, with a corresponding decrease in meter deflection. The ohmmeter scale is calibrated accordingly.

The ac-to-dc conversion circuit of the Heathkit VTVM along with the standard probe is good for ac voltages from 25 Hz to 1 MHz with an accuracy of (±1 dB). However, with an additional rf probe (similar to the Hewlett-Packard), the range can be extended up to 100 MHz. With this probe, the meter is used with the function switch in the dc position, since rectification would now take place in the probe.

The input impedance to the dc bridge measuring circuit in the IMW-11 is 10 M ohms, so that there is greater loading effect caused by this meter in comparison to the Hewlett-Packard model, which has an input impedance of 122 M ohms.

On ac, the input impedance of the Heathkit VTVM is 1 M ohm, and the accuracy is ±5% of full scale. (This resistance is shunted by 35 pF, which limits the high frequency capability.) On dc, the input impedance is 11 M ohms (1 M ohm in probe), and the accuracy is ±3% of full scale.

4-6. FULL-WAVE VOLTAGE DOUBLER

The full-wave voltage doubler does not rely upon the clamping action and subsequent rectifying, as in the half-wave doubler. This can be seen in the circuit of Fig. 4-12(a), where each diode allows one capacitor to be charged up to the peak voltage of the input. The action can perhaps be more clearly seen when the circuit is redrawn in the bridge layout of Fig. 4-12(c). On the positive half-cycle of v_s, D1 conducts and charges C_1 to V_m. (D2 is reverse biased.) On the negative half-cycle, C_2 charges through D2, also to V_m. Since R_L is across the series combination of C_1 and C_2, the total voltage is $2\ V_m$. The exact value of v_o will depend upon how much load resistance R_L is present, the voltage drop across the diodes, and the size of the capacitors.

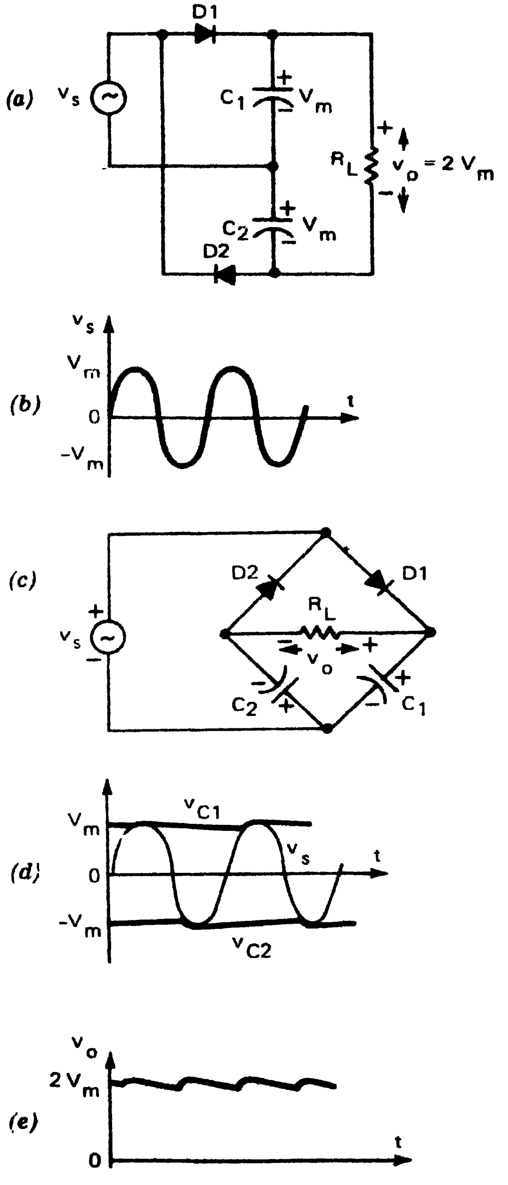


Fig. 4-12. (a) Circuit of a full-wave voltage doubler.
(b) Input waveform, v_s.
(c) Full-wave voltage doubler circuit redrawn in a bridge layout.
(d) Waveforms of v_{C1} and v_{C2}.
(e) Resulting output waveform with $V_{dc} \approx 2\ V_m$.

Since two charging pulses occur every cycle there will be a ripple frequency in the output twice that of the line. (See Fig. 4-12(e).) This is a full-wave rectifier, but the ripple is greater than in a conventional full-wave rectifier and the voltage regulation is worse. However, no transformer is necessary. One minor drawback from a construction and testing standpoint is the fact that there is no common connection between the ground of the input and the ground of the output. This disadvantage is not present in the half-wave voltage doubler.

The circuit of Fig. 4-12 does provide a ready method of obtaining two voltages of opposite polarity with respect to the input ground. This is shown in Fig. 4-13, where resistors have been placed across the capacitors to discharge them after the input voltage is removed.

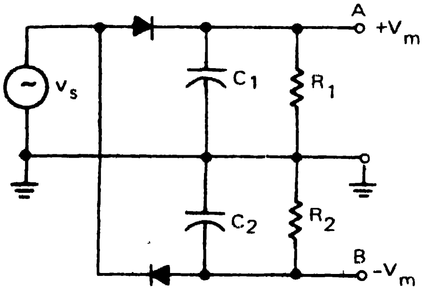


Fig. 4-13. A full-wave voltage doubler circuit to provide opposite polarity output voltages.

4-7. HALF-WAVE VOLTAGE MULTIPLIER

Where the voltage doubler circuit of Fig. 4-10 does not provide a high enough output with the available input, it is possible to obtain any degree of voltage multiplication by the use of circuits of the type shown in Fig. 4-14.

A load connected between X and Y would have a dc voltage of approximately 3 V_m and from W to Z an output of 4 V_m, since capacitor C_1 charges up to the peak of the input voltage, V_m, and the other capacitors to 2 V_m. Note the directions of the diodes in this circuit. They make Y negative with respect to X, and Z negative with respect to W.

Such a circuit (with the diodes alternated) is very frequently used to provide the high accelerating voltage (3 kV - 14 kV) found in cathode-ray tubes in oscilloscopes. This circuit eliminates the bulky and dangerous capacitors and heavily insulated transformer windings that would be necessary with conventional power supplies. In fact, one manufacturer of oscilloscopes reduces the size of the capacitors even further by using an input frequency of 60 kHz instead of the 60 Hz line frequency. In this way, the size of the required filter capacitors is reduced by a factor of 1000, allowing relatively inexpensive disc capacitors (0.02 - 0.03 μF) to be used instead of expensive 20 μF capacitors. The 60 kHz voltage that supplies the primary winding of the high-voltage transformer is developed from a class C oscillator as shown in Fig. 4-15. This circuit employs a voltage tripler, although a septupler (seven times multiplication) may be used with some high voltage supplies. With a 1000 V_{rms} output at the secondary of the transformer, the dc output would be $3\sqrt{2}$V, or approximately 4 kV.

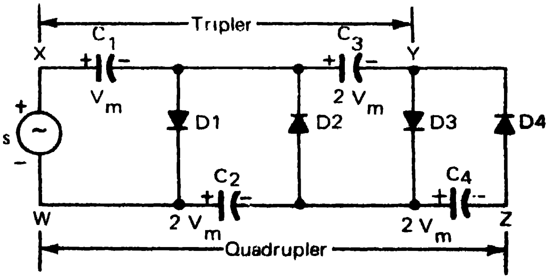


Fig. 4-14. Voltage multiplier showing tripling and quadrupling action.

Although the regulation of a voltage multiplier is usually very poor and its ripple high, the constant low current required by the CRT

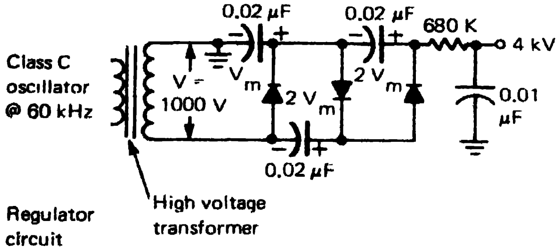


Fig. 4-15. Circuit of a voltage tripler to provide a CRT with its high accelerating voltage.

accelerating anodes provides a light enough load to avoid these problems.

4-8. SUMMARY

(1) The diode clamper is a circuit employing a series-connected capacitor and diode to shift the input voltage by an amount equal to the peak of the input, clamping the output voltage near zero volts.

(2) The output from a diode clamper, also known as a shunt rectifier, is taken across the diode and may be positive or negative, depending upon the diode's polarity.

(3) If an RC filter is added to the output of a shunt rectifier that has a sinusoidal input, the dc output voltage will be $+V_m$ or $-V_m$, with a ripple dependent upon load size.

(4) The shunt rectifier with a filter is used in some VTVMs to convert the ac being measured to a proportional dc voltage that can operate the standard dc indicating meters.

(5) Ac voltmeters employing a shunt rectifier react to the peak of the alternating input waveform and are calibrated to read true rms values only when used with pure sine waves.

(6) A half-wave or cascade voltage doubler is a diode clamper circuit followed by a rectifier and filter, producing a dc output voltage of $-2\ V_m$ or $+2\ V_m$, depending upon diode polarities.

(7) The half-wave voltage doubler may be used to provide a high dc output voltage (typically 3 kV) or to convert ac into dc in a VTVM application.

(8) A full-wave voltage doubler uses two diodes to charge two series-connected capacitors to a total output voltage of $2\ V_m$, with a ripple frequency twice that of the input.

(9) The full-wave voltage doubler, which does not rely upon a clamping and rectifying action, can be used to provide two dc output voltages of opposite polarity with respect to ground and equal to $\pm V_m$.

(10) The half-wave voltage multiplier may be used to provide any degree of voltage multiplication by cascading diodes and capacitors. In this manner accelerating potentials of 14 kV may be obtained for oscilloscope CRT applications.

QUESTIONS

Q4-1. Why is a diode clamper also referred to as a shunt rectifier?

Q4-2. How does the output of a diode clamper with a sinusoidal input differ from the output of a shunt rectifier and filter combination with the same input in:

(a) Waveform?
(b) Dc voltmeter reading?

Q4-3. Sketch a graph showing how the dc output voltage of a shunt rectifier with filter varies with ac input voltage. On the same axes sketch the relation between the dc output voltage of a voltage doubler circuit and its ac input voltage. What is the relation between the slopes of these two lines?

Q4-4. What is the purpose of placing a shunt rectifier and filter in the probe of an ac voltmeter designed to work at very high frequencies, instead of in the housing of the voltmeter?

Q4-5. If a voltage doubler ac-to-dc conversion circuit provides twice the dc output of a shunt rectifier and filter combination, why was the latter chosen to be placed in the probe of some high frequency ac voltmeters?

Q4-6. Explain the following statement: "An ac voltmeter employing a shunt rectifier reacts to the peak of the alternating input waveform, so that the reading represents true rms values for pure sine waves only."

Q4-7. What effect would the shorting out of D2 have on the waveform of v_o in the circuit of Fig. 4-9?

Q4-8. What is the PIV across D2 in Fig. 4-9, in terms of V_m?

Q4-9. Draw the portion of the voltmeter (ac mode) in Fig. 4-11, from the input probe to Y, in schematic form, indicating the shunt rectifier and filter circuits. Why is the shunt rectifier and filter bypassed when the voltmeter is used in the dc mode?

Q4-10. What is the PIV across D1 in Fig. 4-12, in terms of V_m, and what is the ripple frequency in the output when operating at 400 Hz?

Q4-11. Draw a circuit of a half-wave voltage septupler that provides a positive dc voltage of approximately 7 V_m. Show polarities.

PROBLEMS

P4-1. Draw the waveform observed on a oscilloscope in the dc mode when connected between points A and B with ground at B in each of the circuits on page 61 with the indicated input waveforms applied. Indicate voltages and the zero level. Also give the numerical values of the reading of a dc voltmeter connected across A and B, ground at B.

P4-2. What is the PIV across the diodes in P4-1(a), (b), (c)?

P4-3. (a) What is the PIV across each diode in P4-1(d)?

(b) What voltage does each capacitor charge up to?

P4-4. Draw a graph of the dc output voltage, V_o, against input rms voltage, V_s, in Fig. 4-6(b). Make V_s vary from 0 V to 5 V. Neglect the nonlinearity of the diode. What effect would the actual characteristics of a real diode have on the graph? How does this show up on the low end of an ac voltmeter scale that has a full scale deflection of only 1.5 V?

P4-5. (a) What is the PIV across diodes D1 and D2 in Fig. 4-10?

(b) What must be the minimum voltage rating of capacitors C_1 and C_2?

(c) Determine the percent ripple in the output when a 10 M ohm resistive load is connected. [Hint: Use Eq. 3-77, remembering that the ripple frequency is 60 Hz, not 120 Hz, and that C_2, C_3 and the 680 K ohm resistor constitute a π-type CLC filter with the 680 K ohm resistor taking the place of X_{L1}.]

P4-6. Draw a graph showing how the open-circuit dc output voltage from Fig. 4-10 varies with the rms input voltage V (on transformer secondary) for values of V from 0 V to 5 V. (Assume ideal diodes.)

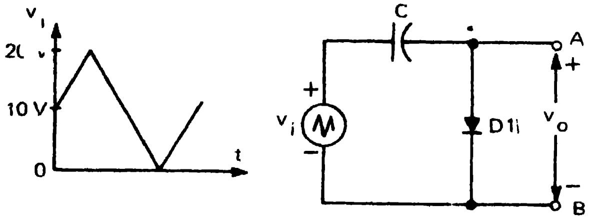
P4-7. Referring to the full-wave voltage doubler circuit in Fig. 4-13:

(a) Determine the rms value of V_s to produce 200 V dc between A and B.

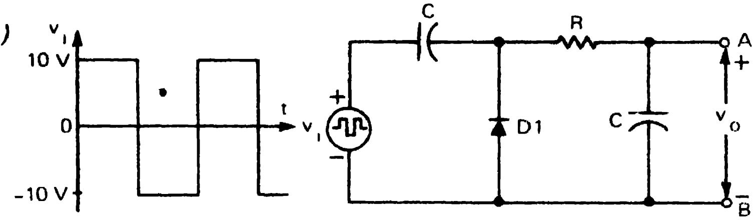
(b) With the input as in (a), $C_1 = C_2 =$ 20 μF, and $R_1 = R_2 = 22$ K, calculate the percent ripple in the output between A and ground with no load connected. [Hint: Use Eq. 3-25 multiplied by 2 (half-wave), with $R_L = R_1$.]

(c) What is the ripple frequency between A and ground, and between A and B?

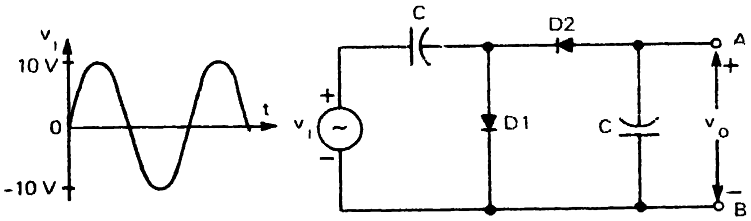
(a)


v_i
20 V
10 V
0
t
C
v_i
+
−
D1
A
+
v_o
−
B

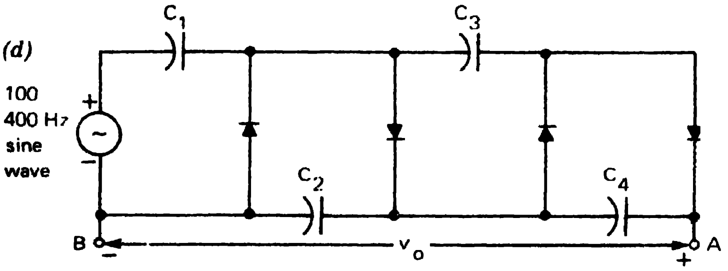
(b)


v_i
10 V
0
-10 V
t
C
R
v_i
+
−
D1
C
A
+
v_o
−
B

(c)


v_i
10 V
0
-10 V
t
C
D2
v_i
+
−
D1
C
A
+
v_o
−
B

(d)


100
400 Hz
sine
wave
+
−
C_1
C_3
C_2
C_4
B
−
v_o
A
+

(d) What is the PIV across the diodes?

(e) What must be the minimum voltage rating of the capacitors?

(f) What is the peak diode current when 10 K ohm loads are connected from A to ground and B to ground? (Eq. 3-52 and 3-54)

P4-8. Design a half-wave voltage quintupler to provide 8 kV. Specify:

(a) The rms value of the 60 Hz input voltage.
(b) The minimum PIV rating of the diodes.
(c) The minimum voltage rating of each capacitor

5

SEMICONDUCTOR PHYSICS OF DIODES

5-0. INTRODUCTION

The principles underlying the operation of the semiconductor diode also apply to transistor operation. The makeup of P-type and N-type materials will be discussed, and an analysis of the diode under conditions of forward and reverse bias will be made. Finally, the effects of temperature variations on PN junctions will be considered.

5-1. REVIEW OF ATOMIC STRUCTURE

The way in which atoms of a material or element combine with each other and behave in the presence of an electric or magnetic field depends upon their atomic structure. Since most semiconductor devices use germanium or silicon as a base material, the following discussion will be limited to these materials.

Consider the diagram representing the Bohr concept of the atomic structure of germanium as shown in Fig. 5-1.

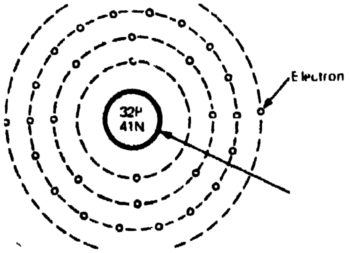


Fig. 5-1. Diagram of the Bohr concept of an atom of germanium.

The nucleus consists of 32 protons, each having a positive charge equal in magnitude to the charge of an electron, and 41 neutrons, each having no charge, just mass. The mass of protons and neutrons is approximately 1850 times as large as that of the electron.

Mass of electron = 9.11×10^{-31} Kg.

Charge of electron = 1.6×10^{-19} coulomb. (Negative)

The four orbits contain 32 electrons in the following manner:

1st Orbit:	2 Electrons		(Full)
2nd Orbit:	8 Electrons	2 Subshells	(Full)
3rd Orbit:	18 Electrons	3 Subshells	(Full)
4th Orbit:	4 Electrons	4 Subshells	(Can Hold 32)

The shells fill up from the center, until a number of electrons equal to the number of protons is reached. The atom as a whole is *intrinsically neutral*. However, the fourth shell which could hold 32 electrons has only 4 electrons in it. This outer shell, often called the *valence* shell determines how the atom will bond or react with other atoms.

For example, in an insulator, the valence shell is completely filled, and all the electrons are tightly bound to the nucleus. The application of an emf has no effect upon these electrons so no current flows. If, however, an exceedingly large potential is applied, it is possible for electrons to be torn from the outer orbit to constitute a current flow. Under this condition the insulator is said to have broken down.

In good conductors, for example copper, the outer shell has only 1 electron in an orbit that can hold 32. This electron is very loosely bound to the nucleus and easily becomes a free-electron at room temperatures, ready to be moved along in the crystalline structure under the influence of an applied electric field. The free-electron acquires sufficient energy conduction band (either by heat or doping) then this reference line or Fermi level would shift upward to indicate this increase in energy. Alternately, a downward shift would indicate a decrease in energy level or a reduction of electrons in the conduction band.

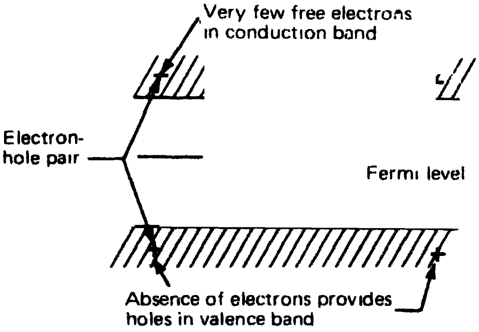

Very few free electrons in conduction band
e
Electron-hole pair
Fermi level
Absence of electrons provides holes in valence band

Fig. 5-5. Diagram of energy levels in pure germanium above 0°K.

The question may arise, why do only certain electrons get enough energy to leave the valence shell and not all of them? The answer is that when an electron breaks a bond, to become free, the remaining electrons exercise a stronger influence upon each other in their attraction to the nucleus, making it more difficult for another electron to leave.

5-5. CONDUCTION IN PURE GERMANIUM

It should be noted that the electron that is free may provide conduction by drifting through the crystal when acted upon by an emf, or it may elect to recombine with a hole recently vacated by another electron and continue to "jump" from hole to hole in the same direction as a free electron in the conduction band. However, the effect is that of a hole or positive charge moving in the opposite direction. Because these are two different mechanisms of current flow, we can refer to the first as "electron flow in the conduction band" and to the second as "hole flow in the valence band," as depicted in Fig. 5-6. In a metal, current flow is by the movement of electrons in the conduction band. However, in semiconductor materials, *both* currents flow at the same time, and each has a different effect upon the characteristics of diodes and transistors. This will be referred to later in the P-type and N-type materials.

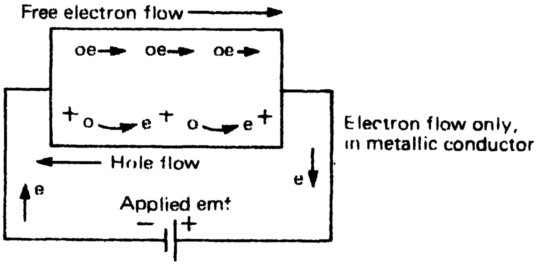


Fig. 5-6. Two types of current flow in pure germanium showing electron flow in the conduction band and hole flow in the valence band.

Although diodes and transistors consist of two distinctly different materials, P-type and N-type, the device is usually made by doping a single crystal of germanium (or silicon) to give the properties desired. However, for ease of discussion, let us consider the P-type and N-type materials separately, then combine them into a diode.

5-6. N-TYPE MATERIAL

The electrical conductivity of pure germanium (or silicon) can be drastically increased (approximately 1000 times) by the addition of an impurity material, in the ratio of 1 atom per 10 million atoms of germanium. If the impurity is pentavalent (5 electrons in its valence shell), such as phosphorus or arsenic, there will be one electron that is not required for bonding when a phosphorus atom is substituted for a germanium atom in the crystal. (See Fig. 5-7.) Since it only requires 0.01 eV to liberate this fifth electron from the phosphorus atom (which is supplied when the material is at room temperature) there will be a large increase in the number of free electrons in the crystal.

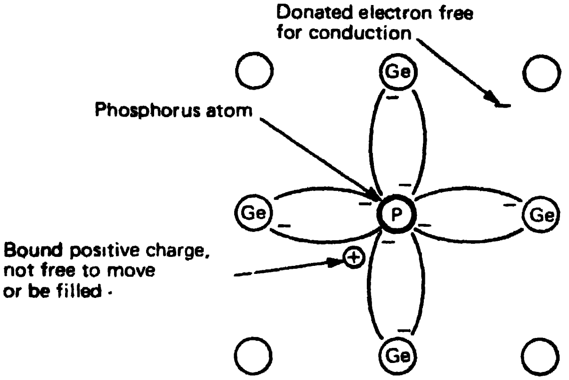

Donated electron free for conduction
Phosphorus atom
Bound positive charge, not free to move or be filled
Ge
Ge
Ge
Ge
P

Fig. 5-7. Portion of a germanium crystal with a phosphorus atom substituted for a germanium atom, to provide an N-type material.

The escaping of the electron from the phosphorus atom leaves a bound positive charge which cannot move. Recombination does not take place because of the shielding effect of the remaining electrons. Since the phosphorus has been responsible for donating electrons to the crystal, it is called a *donor* material. The germanium is said to have been *doped* and is now an *extrinsic* material. (More free electrons than holes.)

As in an intrinsic crystal, thermal agitation still causes hole-electron pairs, but, because of the abundance of the donated electrons, combination of holes and electrons occurs rapidly, so that there are very few holes in an N-type material (even less than in pure germanium). Because the free electron density is predominant (over holes), the material is called N-type (N for negative), even though the crystal as a whole is still neutral.

Current flow through the N-type material is due primarily to the donated free electrons, so that they are called the *majority current carriers*. The holes are termed the *minority current carriers*, since current flow in the valence energy band is very small. It should be noted that the Fermi level has been raised to signify the increase in the number of electrons in the conduction band due to the donor material, as shown in Fig. 5-8.

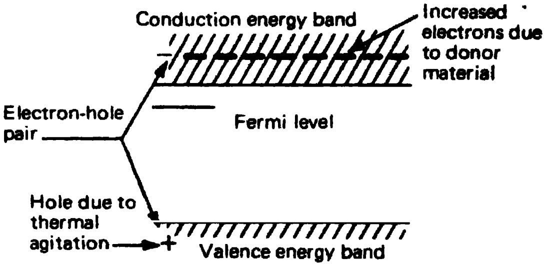

Conduction energy band
Increased electrons due to donor material
Electron-hole pair
Fermi level
Hole due to thermal agitation
Valence energy band

Fig. 5-8. Energy level diagram for N-type semiconductor, showing the raised Fermi level.

5-7. P-TYPE MATERIAL

A corresponding doping effect can be had by adding one atom of an impurity material that is trivalent (3 electrons in its valence shell) for every 10^5 atoms of germanium so that the bond will be incomplete. Materials such as boron, indium, gallium, or aluminum are used and, since the effect is the opposite to donating an electron, they are called *acceptor* materials. Since there is no electron to complete the bond a hole exists, so that the impurity atom can accept an electron from a nearby atom, creating another hole. This is illustrated in Fig. 5-9. At room temperature, again electron-hole pairs are formed, but, because of the greatly increased number of holes, there are very few free electrons (even less than in pure germanium). When the impurity atom accepts an electron for bonding purposes it acquires a bound negative charge. The hole that was created in the adjoining germanium atom is available to be filled by another electron (at the valence energy level) from another atom. Hence the hole moves through the crystal. Since the holes predominate they are called the majority current carriers while the free electrons are the minority current carriers. The holes being positive account for the name P-type. Most current flows through a P-type material by means of electrons filling holes in the valence energy band; this mechanism is quite different from that of the electrons moving in the conduction energy band, which produces an extremely small current in a P-type material. As could be expected, the Fermi level has dropped since there has been a reduction of free electrons in the conduction energy band. See Fig. 5-10.

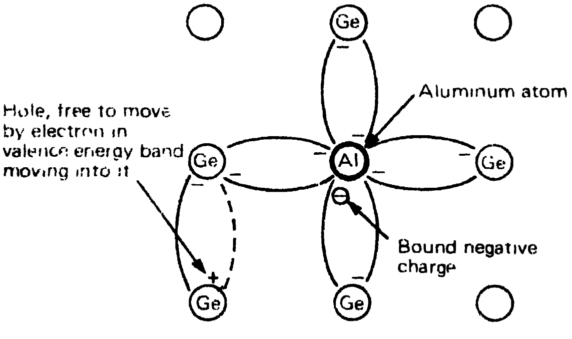

Ge
Aluminum atom
Hole, free to move by electron in valence energy band moving into it
Al
Bound negative charge
+

Fig. 5-9. Portion of a germanium crystal with an aluminum atom substituted for a germanium atom to provide a P-type material.

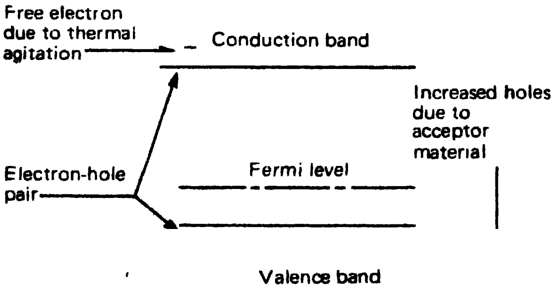

Free electron due to thermal agitation
Conduction band
Increased holes due to acceptor material
Electron-hole pair
Fermi level
Valence band

Fig. 5-10. Energy level diagram for P-type semiconductor.

5-8. CONDUCTION PROCESSES IN N-TYPE AND P-TYPE SEMICONDUCTORS

As a summary of the conducting properties of N-type and P-type materials under the influence of an applied electric field consider the two diagrams in Fig. 5-11.

It is interesting to note that, for the same degree of doping in the N and P types, the greater mobility of the free electrons results in the N-type semiconductor having approximately half the resistivity of the P-type.

5-9. PN JUNCTIONS

PN junction diodes may be of the grown, alloyed, or diffused type. In the grown type, impurities are introduced during the growing of the crystal to give the N- and P-type properties. An alloyed junction is made from an N-type slice of germanium (or silicon) by melting a pellet of trivalent indium placed on the slice. The indium is absorbed into the germanium to produce a P region and hence a PN junction. In the diffusion process the semiconductor wafers of one conductivity (P or N) are placed in a vessel that contains an oxide of the impurity to be added. The combination is passed slowly through a furnace with accurately known and controlled temperatures between 800°C and 1200°C depending upon the type of junction desired. A gas of the impurity atoms diffuses into the semiconductor material and forms a thin layer of opposite conductivity, resulting in a PN junction.

Thus the PN junction is *not* formed by joining two crystals of P- and N-type materials. However, to describe the mechanism that takes place at the time the junction is formed, let us consider placing an N-type and a P-type semiconductor side-by-side to "make" a PN junction. This is represented diagrammatically in Fig. 5-12.

At the moment of "joining together" some conduction electrons will move from the N-type material into the P-type and undergo

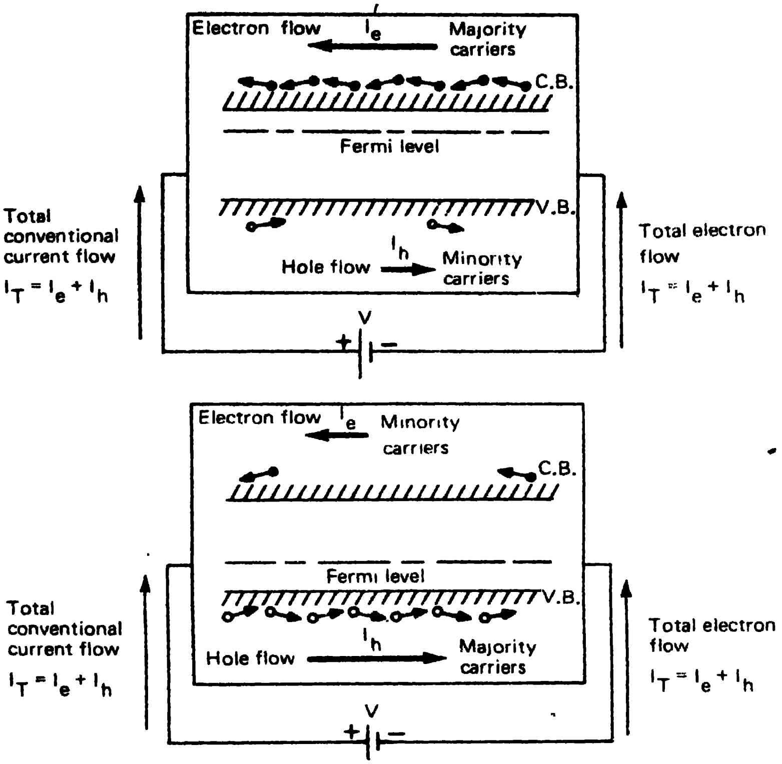

Electron flow
I_e
Majority carriers
C.B.
Fermi level
V.B.
I_h
Hole flow
Minority carriers
V
+
−
Total conventional current flow
$I_T = I_e + I_h$
Total electron flow
$I_T = I_e + I_h$
Electron flow
I_e
Minority carriers
C.B.
Fermi level
V.B.
I_h
Hole flow
Majority carriers
V
+
−
Total conventional current flow
$I_T = I_e + I_h$
Total electron flow
$I_T = I_e + I_h$

Fig. 5-11. (a) Current flow in N-type semiconductor. (b) Current flow in P-type semiconductor.

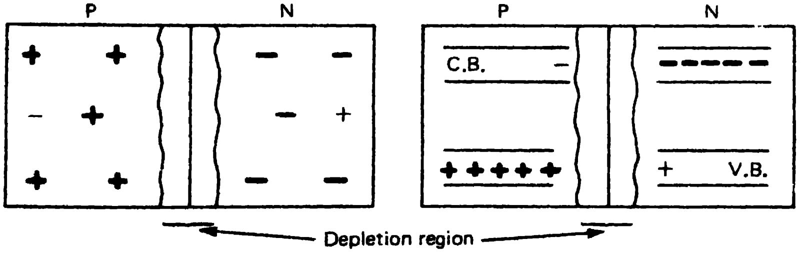

P
N
P
N
C.B.
V.B.
Depletion region

Fig. 5-12. Diagram showing the distribution of electrons and holes in a PN junction, and the resulting depletion region.

electron-hole recombination with the holes in the valence band of the P-type. Thus holes will flow from the P-type to the N-type. This only continues for a very short time since the movement of holes and electrons will cause the P-type to become negatively charged (since it has acquired some electrons) and the N-type positively charged (since it has lost some electrons). This process of establishing equilibrium causes a *depletion region* at the junction where there are neither holes nor conduction electrons, as shown in Fig. 5-12. No further movement will occur since there is effectively a potential developed across the junction of a few tenths of a volt. The P-type being more negative than the N-type will prevent further electron incursion, while the N-type is positive with respect to the P-type and will oppose the flow of holes.

The depletion region, sometimes called the transition region, is in the order of 0.0005 mm, but this varies as the bias across the junction is varied. In many diodes, the method of construction makes the P-type more heavily doped, so that the depletion region is mostly in the N-type material.

The potential profile that exists in the PN junction is drawn in Fig. 5-13.

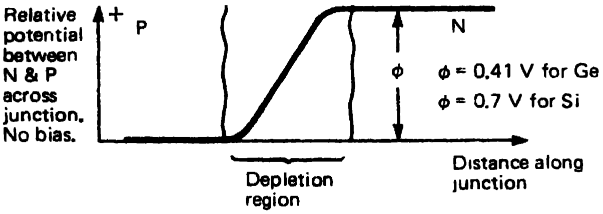


Fig. 5-13. Potential profile across the depletion region in a PN junction.

The potential difference could be represented by a battery across the junction as illustrated in Fig. 5-14.

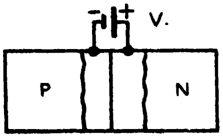


Fig. 5-14. Diagram showing how the potential across an unbiased junction may be represented by a battery of value V_j.

(V_j can be calculated but not measured due to the contact potential between the semiconductor and the metallic measuring instrument.) This can be thought of as a potential hill or barrier that can be increased or decreased by the application of external voltages.

5-10. FORWARD-BIASED JUNCTION

Consider the application of an electric field to a PN junction as shown in Fig. 5-15(a).

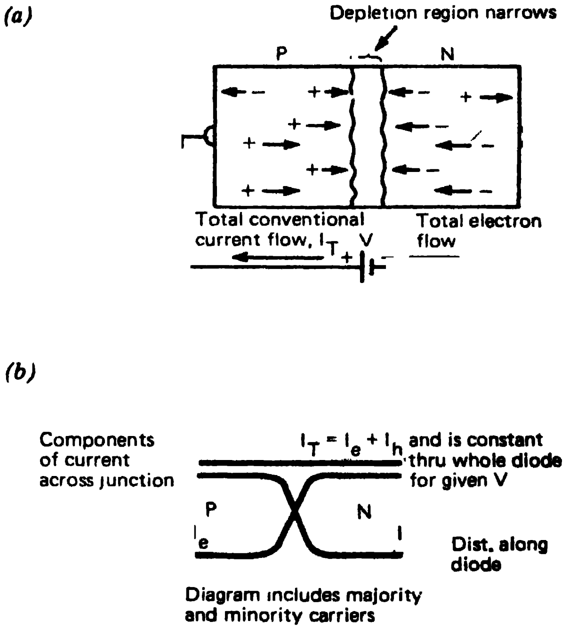

(a)
Depletion region narrows
P
N
Total conventional current flow, I_T
V
+
−
Total electron flow
(b)
Components of current across junction
$I_T = I_e + I_h$ and is constant thru whole diode for given V
P
N
I_e
I
Dist. along diode
Diagram includes majority and minority carriers

Fig. 5-15. (a) Forward-biased PN junction showing motion of electrons and holes. (b) Graph showing components of current across the PN junction.

The potential across the junction will be lowered to a value Φ − V, allowing the positive side of V to repel holes in the P-type material and the negative side of V to repel electrons in the N-type material. The flow of current across the junction itself is due to two components; (a) electron flow in the conduction band (which decreases the further that electrons flow into the P material), and (b) hole flow in the valence band (which also decreases the further that holes flow into the N material). Hole flow and electron flow are approximately equal in the center of the transition region, as shown in Fig. 5-15(b). In addition to this large current of the majority carriers, there is a small current due to the minority carriers, but it is almost negligible. Note: Forward bias means positive to P, negative to N. (See Fig. 5-16.)

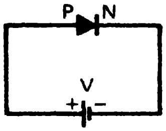


Fig. 5-16. Forward-biased PN junction showing the diode symbol.

5-11. REVERSE-BIASED JUNCTION

Consider reversing the applied voltage as in Fig. 5-17.

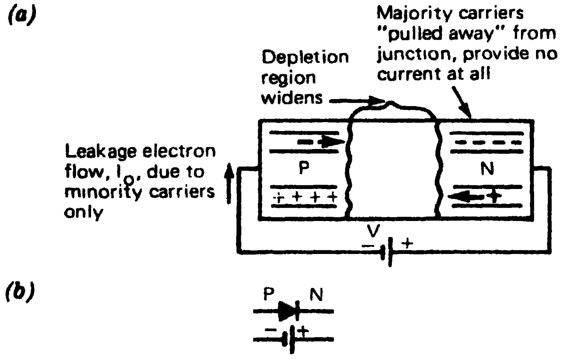

(a)
Majority carriers "pulled away" from junction, provide no current at all
Depletion region widens
Leakage electron flow, I₀, due to minority carriers only
P
N
V
−
+
(b)
P
N
−
+

Fig. 5-17. (a) Reverse-biased junction. (b) Reverse bias showing diode symbol.

The effect is to increase the potential hill to $\Phi + V$, preventing the flow of majority carriers. The depletion region widens due to the positive terminal of the battery drawing electrons from the N-type and the negative terminal attracting holes from the P-type. At first it seems as if no current will flow at all due to no carriers crossing the junction. However, the minority carriers are unaffected by this barrier so that the free electrons in the conduction band of the P-type are repelled by V and cross the junction, as do the holes in the N-type, both types of carrier constituting a small reverse current flow, I_o. Since the density of the minority carriers depends upon the amount of thermal agitation, this reverse current is very temperature dependent. For the same reason, the forward current is also temperature dependent, but to a much smaller degree, as shown by the next section.

5-12. EFFECT OF TEMPERATURE ON PN JUNCTION DIODES

From a consideration of the energies of the carriers crossing the depletion region, an involved calculation yields the following relation between diode current, voltage and temperature:*

$$I = I_o \times (e^{\frac{V}{\eta V_T}} - 1), \qquad (5\text{-}1)$$

where I = the diode current, forward if positive, reverse if negative, same units as I_o.

= the diode reverse current, also called the reverse saturation current, at temperature T.

= the diode voltage, positive for forward bias, negative for reverse bias, in volts.

η = 1 for germanium, 2 for silicon.

$V_T = \frac{T}{11{,}600}$, a quantity in volts, dependent upon temperature.

---

*Millman, J., and Halkias, C., *Integrated Electronics*, McGraw-Hill, New York, 1972, Section 3-3.

T = the temperature of the diode junction, °K = °C + 273°.

At room temperature, 22° C (72° F), $T = 295°K$ and $V_T = 0.025$ volt. Thus, for germanium,

$$I = I_o(e^{40V} - 1) \tag{5-2}$$

and for silicon,

$$I = I_o(e^{20V} - 1) \tag{5-3}$$

where I_o is the reverse saturation current at room temperature. These equations indicate that when the diode is forward biased (V positive) the current increases exponentially with voltage above a few tenths of a volt, 0.2 V for germanium and 0.6 V for silicon.

When V is negative for reverse-biased conditions $I = -I_o$. That is, when the reverse bias exceeds approximately 0.1 volt, the reverse current saturates at I_o and is theoretically constant at this value. In actual practice many silicon diodes show an increasing amount of reverse current with increase in reverse bias. This is due to surface imperfections in the diode's depletion region permitting an additional leakage current to occur.

As mentioned earlier the reverse current I_o is very temperature dependent and may be shown to increase 7%/°C for both germanium and silicon. This represents a doubling for every 10°C rise. The reason that germanium *seems* more temperature dependent is that its reverse current is approximately 1000 times larger than the silicon reverse current, so that it is a quantity that is not insignificant.

For forward-bias operation the change in the voltage required for any given value of current is -2.5 mV/°C for either germanium or silicon. Conversely, the current for a given value of voltage will increase with temperature.

Actually, if the equation for I is examined, an increase in T (and V_T) will serve to *reduce* the exponent of e, tending to *reduce* I. However, the accompanying *increase* in I_o will more than make up for this reduction with a resulting overall increase in I, as shown by the following example.

Example 5-1. Determine the forward current for a germanium diode at room temperature when the voltage across it is 0.3 volt, and compare with the current when the temperature rises 50°C. Leave the answer in terms of I_o.

Solution. At room temperature, for germanium,

$$\begin{aligned} I_{22^\circ C} &= I_o(e^{40V} - 1) \\ &= I_o(e^{12} - 1) \approx I_o \cdot e^{12} = I_o\, e^6 \cdot e^6 \\ &= I_o \times 400 \times 400 = 160{,}000\, I_o. \end{aligned}$$

$$\text{At } 72^\circ\text{C, } V_T = \frac{T}{11{,}600} = \frac{345}{11{,}600} = 0.0298 \approx 0.03 \text{ volt.}$$

Due to a doubling of I_o for every 10°C rise, the value of reverse current at 72°C, I_o', will be $I_o \times 2^5 = 32\, I_o$.

(1 → 2 → 4 → 8 → 16 → 32)

$$\begin{aligned} \therefore I_{72^\circ C} &= I_o'(e^{\frac{V}{V_T}} - 1) \\ &= 32I_o(e^{\frac{V}{0.03}} - 1) \approx 32I_o(e^{33V} - 1) \\ &= 32I_o(e^{10} - 1) \approx 32I_o \times 22{,}000 \\ &= 704{,}000 I_o. \end{aligned}$$

Thus the forward current increases by a factor of approximately 4, which is due to the increase in the reverse saturation current by a factor of 32.

The above example should indicate a method of determining I_o by experimental means for a given diode, especially a silicon diode, whose small value of reverse current is not easily directly measurable. A graph showing the variation in a silicon diode's characteristics with temperature is given in Fig. 5-18. A germanium diode's characteristics are very similar except for the much larger initial value of I_o, which is in the microampere range, and a smaller forward-voltage drop, with a cut-in value of 0.2 V at 25°C.

It should be clear that the reason for the cut-in voltage in a semiconductor diode is the potential barrier, Φ, referred to in Fig. 5-13. The application of an external voltage V reduces Φ, allowing an exponential increase in forward current. However, the barrier is never reduced completely to zero, since the bulk resistance of the diode (in the order of a few ohms) along with the relatively high current flow does not allow the externally applied voltage V to be impressed directly across the junction. Since the ohmic resistance of the diode limits current flow, the diode's characteristics approximate a straight line at high values of forward current.

5-13. MAXIMUM TEMPERATURE OPERATION

The logical question that may be asked is: To what maximum temperature can a diode be subjected? This depends upon the construction and mounting of the diode, but the maximum temperature a germanium *junction* can withstand without damage is in the range of 60 to 100°C and for a silicon junction it is between 150 and 225°C. A germanium seldom is used higher than 75°C, and a silicon seldom

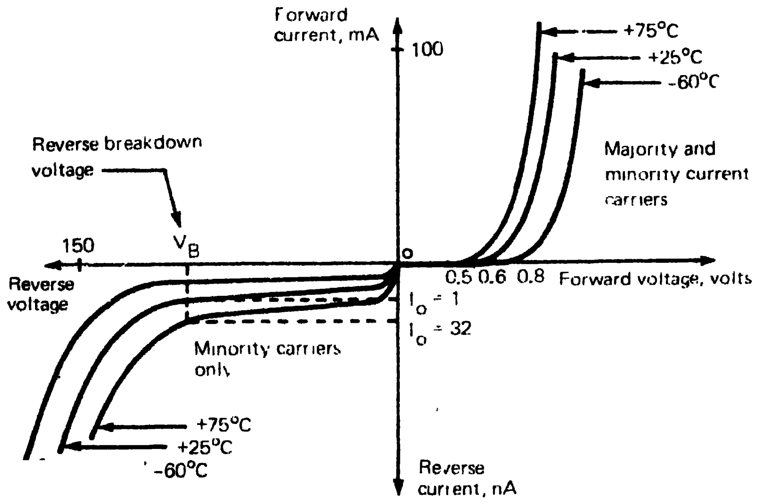


Fig. 5-18. Graph showing the variation in a practical silicon diode's characteristics with temperature change (The change in reverse current from 25°C has not been drawn to scale.)

higher than 175°C. Both have a lower limit of -65°C.

However, a diode rating usually is at a given temperature, which may be the ambient temperature for lower-powered devices, such as 3 amperes half-wave average at 50°C ambient; or it may be the case temperature of the diode for high power rectifiers, such as 500 amperes average at 150°C case temperature. The case temperature is important where heat sinks may be employed to conduct away the heat.

If the temperature of operation is at a higher value than specified, the devices must be derated. The topics of heat sinks and derating will be more fully covered when transistors at elevated temperatures are considered in a later chapter.

For most power rectifying applications silicon diodes are used exclusively, for their low reverse current and high temperature capabilities. Germanium diodes are sometimes found in computer circuit applications where a small forward-voltage drop may be desirable.

5-14. REVERSE BREAKDOWN VOLTAGE

The characteristic curves drawn in Fig. 5-18 show that when the reverse voltage exceeds some value, V_B, the reverse current increases substantially, and the diode is said to be in the breakdown mode, since it is no longer blocking current.

The explanation of breakdown is found in the avalanche effect, in which the minority carriers (which constitute reverse current) cross the depleted region and gain enough kinetic energy to knock bound electrons out of the covalent bonds. These electrons will in turn collide with other atoms and will increase the number of electrons and holes available for conduction. This multiplication effect of free carriers may be represented by the following equation:*

$$M = \frac{1}{1 - \left(\frac{V}{V_B}\right)^n} \qquad (5\text{-}4)$$

where

M is the ratio of carriers leaving the depletion region to the number entering the region,

V is the applied reverse voltage (positive),

V_B is the reverse breakdown voltage,

n is approximately 3 for a silicon junction.

As V approaches V_B, M increases very rapidly, being very small for $V = 0.9\ V_B$ but very large for $V > 0.9\ V_B$. The resulting reverse current is given by:

$$I_R = MI_o, \qquad (5\text{-}5)$$

where I_o is the reverse saturation current before breakdown.

The value of breakdown voltage V_B depends upon the degree of doping in the diode. The breakdown voltage for silicon diodes can be made to occur at a voltage as low as 5 volts with 10^{17} impurity atoms per cubic cm or as high as 1000 volts when doped to a level of only 10^{14} impurity atoms per cubic cm.

5-15. CAPACITIVE EFFECTS IN A PN JUNCTION DIODE

Consider a PN diode with dimensions as shown in Fig. 5-19.

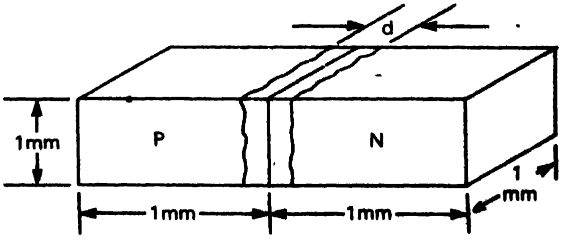


Fig. 5-19. Diagram showing the dimensions of a PN diode.

---

*Gibbons, J., *Semiconductor Electronics*, McGraw-Hill, New York, 1966, Section 6.5.1.

Space-Charge Capacitance

With no applied voltage the width d of the depletion region can be calculated to be 5 X 10^{-5} cm = 5 X 10^{-4} mm = 0.0005 mm and is mostly in the N-type material.* The movement of majority holes and electrons across the junction causes opposite charges to be stored at this distance 'd' apart. This is effectively a parallel plate capacitor whose capacitance C_T (often called space-charge capacitance) can be calculated to be approximately 20 pF with no external bias.

As forward bias is applied (5 mA current) the depletion region decreases to approximately 4 X 10^{-4} mm, and the capacitance C_T increases to 25 pF. Under reverse-bias conditions, (-4V), the depletion region increases to approximately 12 X 10^{-4} mm, with a corresponding decrease of C_T to 8 pF. The variation in space-charge capacitance (sometimes called transition capacitance) with applied reverse voltage is shown in Fig. 5-20 for a TIV 308 silicon voltage-variable-capacitance diode, used in an automatic frequency control (AFC) circuit in an FM tuner.

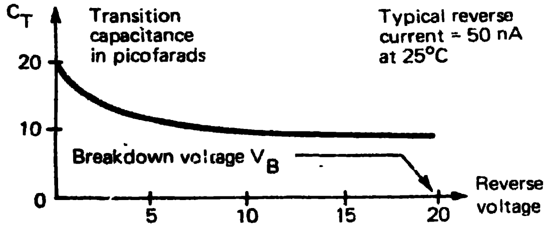


Fig. 5-20. Graph showing variation in capacitance with voltage for a TIV 308 diode.

The application of this property, voltage variable capacitance, appears in varactors, varicaps, or voltacaps. If a diode is used in the reverse-biased manner as a capacitor in an LC resonant circuit, the value of C and hence the resonant frequency can be varied by altering the reverse bias applied to the diode; no moving parts required.

The symbol used for these diodes when used in this manner is as shown in Fig. 5-21 with the equivalent circuit.

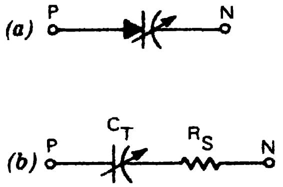


Fig. 5-21. (a) Symbol for voltage-variable-capacitor diode. (b) Equivalent circuit for the diode under reverse bias.

R_S is the body resistance, approximately 1-5 ohms. C_T can be as much as 300 pF in some low voltage Zener diodes.

Diffusion Capacitance

The reason that the capacitance was not shown in Fig. 5-20 for the forward-biased condition is the presence of another capacitance, C_D.

This can be defined by the equation, $C_D = \frac{dq}{dv}$, where dq represents the change in the number of minority carriers stored outside the depletion region when a change in voltage across the diode, dv, is supplied. This is very large in the forward-biased direction, ranging from 8000 pF to as much as 20 μF, certainly swamping out any effect due to C_T.

In the reverse direction, however, C_D is smaller than C_T, and C_T predominates. The large value of C_D must be considered along with the extremely small dynamic forward

---

*Gibbons, J., *Semiconductor Electronics*, McGraw-Hill, New York, 1966, Section 5.6.

resistance r of the diode. The time constant rC_D is so small that it will only be important to consider it when a diode is being used as a switching device at very fast rates. Then, both C_T and C_D have an effect upon the switching capability of the diode. This will be considered at a later time.

5-16. DYNAMIC TESTING OF A DIODE

The dynamic V-I characteristics of a diode may be obtained using commercially available curve tracers or may be displayed on an oscilloscope using the circuit shown in Fig. 5-22, which sweeps the diode voltage at a 60 cycles per second rate. This removes the need for tedious point-by-point plotting and also shows the capacitive effects of the PN junction referred to in the previous section.

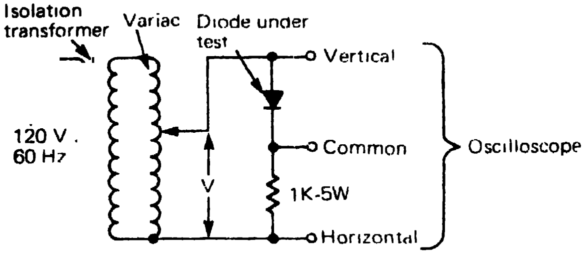


Fig. 5-22. Circuit to obtain dynamic V-I characteristics of a diode.

Except on the more expensive oscilloscopes it is not usually possible to obtain the characteristic curves in the correct quadrants (first and third — I vertically, V horizontally). However, the above circuit does have the advantage of providing the value of forward-voltage drop and reverse breakdown voltage directly, by using the calibrated vertical amplifier on the oscilloscope. (The horizontal is not usually calibrated except on X-Y oscilloscopes.)

The trace obtained on an oscilloscope for a 1N2615 diode is shown in Fig. 5-23. (The quadrants in which the trace will appear will depend upon the particular oscilloscope used.)

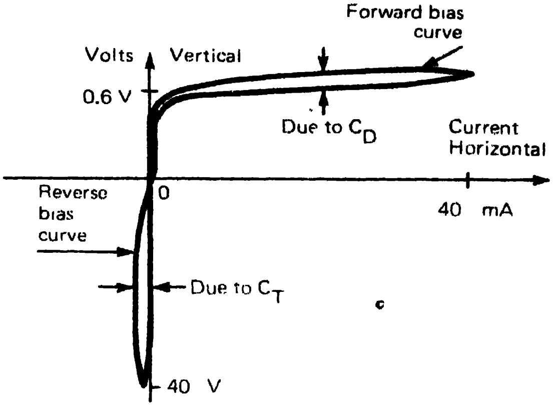


Fig. 5-23. Trace obtained on an oscilloscope for a 1N2615 silicon diode, showing the splitting of the trace due to capacitive effects.

Note: Different scales were used to observe forward and reverse characteristics.

To obtain the reverse characteristic, maximum horizontal sensitivity was used to show the splitting of the trace. This indicated the presence of the transition capacitance C_T of the diode, which was determined to be approximately 0.005 μF (5000 pF) by adding this value of capacitance externally across the diode and noting that the width of the loop doubled in the center portion of the reverse trace. (There was no effect on the forward curve.) Similarly, using maximum vertical sensitivity on the oscilloscope, it was observed that the addition of a 1 μF capacitor across the diode almost doubled the width of the loop in the forward direction, indicating the much larger value of the diffusion capacitance, C_D. For these measurements V in Fig. 5-22 was adjusted for 30 V. By making measurements for different values of V the variation in C_T with V may be obtained. The reason for the large values of C_T (compared with those in Section 5-15) lies in the fact that the

1N2615 diode is a relatively large diode, whose average current capability is 750 mA. It therefore has quite a large cross-sectional area, which contributes to an increase in capacitance. Its reverse breakdown voltage is in the order of 600 V, so it is not possible to determine V_B with the circuit of Fig. 5-22. If a higher voltage is used to display V_B, then a higher value of series resistance must be used to limit the current that would flow in the forward direction.

5-17. SUMMARY

(1) A semiconductor atom of pure germanium or silicon contains four electrons in its valence shell and is closer to a conductor in electrical conductivity.

(2) Semiconductors have a negative temperature coefficient of resistance; that is, current flow increases with an increase in temperature.

(3) Conduction in a pure semiconductor consists of electron flow in the conduction band and hole flow in the valence band.

(4) An N-type material is formed by adding an impurity material such as phosphorus that is pentavalent (donor), causing an increase in free electrons.

(5) A P-type material is formed by adding an impurity material such as indium that is trivalent (acceptor), causing an increase in holes.

(6) The majority current carriers in an N-type material are electrons, while the minority carriers are holes. The reverse is true for a P-type material.

(7) A PN junction is formed in a process which may be of the grown, alloyed, or diffused type.

(8) A PN junction has a built-in potential hill or barrier that exists across the depletion region and may be increased or decreased by the application of an external voltage.

(9) A forward-biased PN junction has the positive terminal applied to the P side and the negative to the N side. The reverse is true for a reverse-biased junction.

(10) The reverse saturation current of a diode increases approximately 7%/°C, while the change in the forward voltage characteristic for a given value of current is -2.5 mV/°C, for either germanium or silicon.

(11) A germanium device is seldom used at a temperature higher than 75°C, a silicon seldom higher than 175°C.

(12) A PN junction exhibits a transition capacitive effect of several picofarads in reverse bias, but may have many microfarads of diffusion capacitance in the forward-biased state.

QUESTIONS

Q5-1. Draw a diagram representing the Bohr concept of the atomic structure of silicon, given that the atomic number (number of electrons) is 14. Indicate the electrons in their orbits, the number of protons and the neutrons if the atomic weight of silicon is 28.

Q5-2. How many electrons are in the valence shell of:

(a) germanium?
(b) silicon?

Q5-3. When electrons in a semiconductor acquire sufficient energy to leave the valence energy band, what are they called and in what energy band are they?

Q5-4. What do you understand by the following terms:

(a) intrinsically neutral?
(b) forbidden band?

Q5-5. Explain why an increase in temperature increases conductivity in a semiconductor, but increases resistivity in most metallic conductors.

Q5-6. The third shell of the atomic structure of silicon consists of 3 subshells. The first can hold 2 electrons, the second 6 electrons and the third 10 electrons. Using this information and the fact that silicon has a total of 14 electrons explain how the covalent bonding of atoms of silicon to form a crystal of silicon is very similar to that for germanium.

Q5-7. Which material, at room temperature, do you think has the higher resistivity, germanium or silicon? What information do you have to support this? How would their conductivities compare at absolute zero °K?

Q5-8. Why do only *certain* electrons in pure intrinsic semiconductors leave the valence shell at elevated temperatures to provide electron-hole pairs, but not *all* the electrons?

Q5-9. Explain the terms "electron flow in the conduction band" and "hole flow in the valence band." Can both types of current flow take place at the same time in a semiconductor?

Q5-10. (a) What types of materials are used for N-type doping?

(b) What is the significance of their pentavalence?

(c) Why is it appropriate to call these doping materials donors?

Q5-11. When a pentavalent phosphorus atom replaces a silicon atom at room temperature what is the effect on:

(a) The number of free electrons in the material?

(b) The number of electron-hole pairs?

Q5-12. For the condition in Q5-11, why does not the liberation of the fifth electron from the phosphorus atom create as many holes in the valence band as it does electrons in the conduction band?

Q5-13. (a) What are the majority carriers in N-type doped silicon?

(b) Is the material still neutral from a charge standpoint?

(c) What constitutes the minority carriers?

Q5-14. (a) Describe the effect of adding a trivalent gallium impurity to silicon.

(b) How does this affect the electron-hole pairs?

(c) What are the majority and minority carriers?

(d) What other materials are used as acceptors?

Q5-15. Summarize and contrast the different mechanisms at work as far as current flow is concerned when a voltage is applied to N-type and P-type materials.

Q5-16. When a PN junction is formed, a depletion region is established at the junction. What is the depletion region and what prevents the *total* recombination of holes and electrons from occurring?

Q5-17. (a) Which PN junction, silicon or germanium, has the higher potential hill or barrier to the flow of majority carriers?

(b) How does this show up in the characteristic curves of silicon and germanium PN junction diodes?

Q5-18. What components of current are there across a forward-biased PN junction?

Q5-19. What happens to the depletion region under conditions of:

(a) forward bias?

(b) reverse bias?

Q5-20. Explain how there can be *any* current flow across a reverse-biased PN junction and why this current is very temperature dependent.

Q5-21. A germanium diode is often said to be more temperature dependent than a silicon diode. Refute this statement with reference to both forward and reverse operation, indicating what might give rise to this statement.

Q5-22. (a) If the forward-voltage drop across a diode is held constant, what will happen to the current through the diode if the temperature is increased?

(b) Which is more sensitive to a change in temperature, forward or reverse current? Why?

Q5-23. What are the approximate upper operating temperatures for germanium and silicon diodes?

Q5-24. Explain in your own words the avalanche effect that accounts for the reverse breakdown voltage (PIV) of a diode.

Q5-25. What does the reverse breakdown voltage of a PN junction diode depend on, and how may it be increased?

Q5-26. (a) What is the effect on the capacitance of a PN junction diode as forward then reverse bias is applied?

(b) What does the capacitive effect arise from in a reverse-biased junction, and what is the capacitance called?

(c) What name is given to diodes that are used for their variable capacitance effect, and where are they applied?

Q5-27. For the circuit in Fig. 5-22, what is the rate at which the current vs. voltage waveform is repeated on the oscilloscope?

Q5-28. Describe the procedure you would use to determine a diode's transition and diffusion capacitances with the aid of the dynamic testing circuit in Fig. 5-22 and an oscilloscope.

PROBLEMS

P5-1. A 1N2615 silicon diode has a reverse current, I_o, of 10 μA at room temperature, 22°C.

(a) Calculate the forward current flowing through the diode at room temperature for the following forward-bias voltages (neglecting diode bulk resistance): 0.3 V, 0.4 V, 0.5 V, 0.55 V, 0.6 V.

(b) Plot the diode current in mA vertically against voltage horizontally.

P5-2. If a 1N2615 silicon diode has a maximum dc current rating of 750 mA, determine the maximum safe forward voltage it may be subjected to without overheating, in the following two ways:

(a) Use the graph from P5-1(b).

(b) Use Eq. 5-3 to solve for V, given $I_o = 10\ \mu$A. Calculate the power dissipated in the diode at this value of current.

P5-3. Calculate the reverse current of a silicon diode at room temperature if a forward voltage of 0.5 V causes 2 A of current.

P5-4. A silicon diode has a reverse saturation current of 1 μA at room temperature, 22°C. Determine the following:

(a) The forward current when the forward bias is 0.56 V at 22°C.

(b) The new value of reverse current when the diode operates at 72°C.

(c) The forward current at 72°C and 0.56 V forward bias.

(d) The reverse and forward current at 2°C and 0.56 V forward bias.

P5-5. A silicon diode is operating at 25°C with a forward bias of 0.6 V and a current of 0.5 A. If the current is held constant at 0.5 A, what voltage will exist across the diode at the following temperatures:

(a) 75°C?

(b) −45°C?

P5-6. Plot a graph of M vs. V/V_B using Eq. 5-4 for a silicon diode, for values of V/V_B from 0 to 0.98. What is M when $V/V_B = 1$, and what is the interpretation of this?

P5-7. A silicon diode has a reverse breakdown voltage, V_B, of 100 V. Its reverse current, I_R, is 1 μA at 22°C when the reverse voltage applied across the diode is 95 V.

(a) What is the diode's reverse saturation current, I_o, at 22°C?
(b) How much current will flow through the diode when it is forward biased with 0.5 V at room temperature?
(c) What will be the forward current at 0.5 V forward bias when the temperature *rises* 30°C?

P5-8. The capacitance of a PN junction diode (like that of a parallel plate capacitor) may be calculated using

$$C = \frac{\epsilon A}{d}$$

where:

ϵ is the permittivity = 8.85×10^{-12} F/m.

A is the cross sectional area in m^2.

d is the width of the depletion region in m.

C is the capacitance in farads.

(a) Determine the space-charge or transition capacitance of the diode shown in Fig. 5-19 under the following conditions:

(i) At zero bias when the depletion region is 0.5 micron (1 micron = 1×10^{-6} meter).

(ii) When a reverse bias of -4 V increases the depletion region to 1 micron.

(b) Determine the effective charge on each side of the depletion region for the condition in (ii).

P5-9. A tuned parallel resonant circuit consists of a 49 μH coil, a 100 pF capacitor and a TIV 308 varactor diode, whose characteristics are shown in Fig. 5-20. Determine the percent change in the resonant frequency of the circuit when the diode's reverse voltage is changed from 1 V to 10 V.

P5-10. (a) In the circuit of Fig. 5-22, what will be the peak current through the diode when V is adjusted to 50 V?

(b) Under the conditions in (a), what will be the maximum reverse voltage across the diode?
(c) What will be the average dissipation in the resistor?
(d) What will the resistor dissipation increase to if V is changed to 100 V?

6

ZENER DIODES, TUNNEL DIODES, AND THERMISTORS

6-0. INTRODUCTION

In the previous chapters we were concerned with using the forward characteristics of a diode, making sure that any reverse voltages that were encountered did not exceed the diode's PIV. In this chapter, we examine a group of diodes referred to as avalanche, breakdown, or Zener diodes, which normally operate in the reverse mode, where their reverse breakdown voltage is used to advantage. Then a Zener diode is used in the design of a voltage regulating circuit and a tube version is also considered. Finally, we look at the characteristics of a tunnel diode and the ac voltage regulating capability of thermistors.

6-1. ZENER DIODES

The current that flows under reverse breakdown is not necessarily destructive if the amount of current is limited to a value such that the diode does not become overheated. By suitably controlling the amount of impurity and making the PN junction free from surface imperfections, the voltage at which breakdown takes place can be made to occur very distinctly and sharply, as shown in Fig. 6-1(a). This type of diode is called a Zener diode and is represented by the symbol shown in Fig. 6-1(b). Although the Zener diode is used in the reverse-biased condition, the symbol is the same as for rectifying diodes. That is, it is normally operated with the anode negative and the cathode positive to provide the characteristics shown in the third quadrant of Fig. 6-1(a).

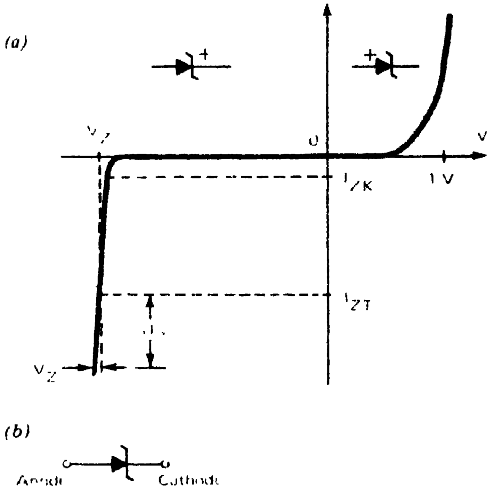


Fig. 6-1. (a) Zener diode characteristics showing the sharp reverse breakdown. (b) Symbol for the Zener diode.

Here it is seen that, once the diode has broken down, very little change in voltage across the diode results from a change in current through the diode. Thus, a Zener diode of known breakdown voltage can be used as a reference voltage when operated in the reverse mode and can be used to regulate voltages.

6-2. ZENER DIODE SPECIFICATIONS

Zener Voltages

The manufacturers specify the value of breakdown voltage known as the Zener voltage, V_Z, at some value of test current, I_{ZT}. This is on the linear portion of the reverse characteristic and corresponds to approximately

one-quarter of the maximum power dissipation capability of the diode. Values of V_Z are available at various values from 2.4 to 200 V with accuracies between 5 and 20%, depending upon cost.

Power Dissipation

Power dissipation in the diode is the product of V_Z and reverse current, I_Z, with maximum power ratings ranging from 150 mW to 50 W.

Breakover Current

Since there is some curvature of the reverse characteristic at low values of I_Z, there may be specified some value of current, I_{ZK}, in the neighborhood of the breakover knee, where the voltage across the diode starts to differ greatly from V_Z.

Dynamic Impedance

There is another specification provided by the manufacturer, which helps to describe how "vertical" is the reverse characteristic shown in Fig. 6-1(a). This is called the Zener dynamic impedance, Z_Z, and is defined as:

$$Z_Z \quad \frac{\Delta V_Z}{\Delta I_Z} \qquad (6\text{-}1)$$

This impedance is usually evaluated in the region of the specified test current, I_{ZT}, and is the reciprocal of the slope of the reverse characteristic. Ideally, Z_Z is zero for a perfectly vertical breakdown curve, but in practice may vary from several ohms to several hundred ohms, depending upon the particular Zener diode voltage and the operating current, as shown in Fig. 6-2. Actually, the equation is more useful in the form:

$$\Delta V_Z = Z_Z \, \Delta I_Z, \qquad (6\text{-}2)$$

since it gives the change in voltage that will occur across the diode due to a change in current through the diode.

For example, a 12 V Zener diode operating at 5 mA has a Z_Z of approximately 25 ohms (from Fig. 6-2) and will have a change of 25 mV across it if the current changes by 1 mA, while at 40 mA operation the same diode will have only 5 mV change for the same change in current. Evidently a Zener diode operated at a higher value of reverse current will have a more constant voltage across it in the face of any changes in diode current. This important property will now be illustrated in the design of a voltage regulator circuit.

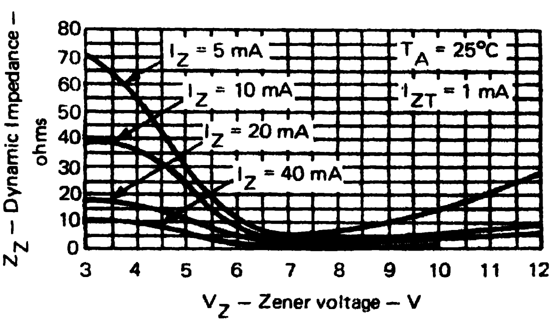


Fig. 6-2. Typical dynamic impedance vs. Zener voltage as a function of Zener current. (Courtesy of Texas Instruments, Incorporated.)

6-3. THE VOLTAGE REGULATOR CIRCUIT

The way in which a Zener diode is used to regulate the voltage across a load in the face of supply voltage or load current variations is shown in Fig. 6-3.

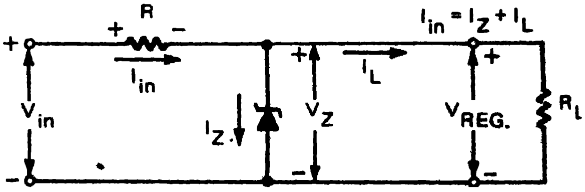


Fig. 6-3. Circuit of a Zener diode voltage regulator showing the series dropping resistor, R.

Note that the reverse characteristics of the Zener diode are being used. The symbol still points in the direction of conventional current flow in the forward-biased connection, even though the diode is not used in this direction. The diode is selected with a V_Z equal to the voltage desired across the load. The Zener diode acts as a "bypass valve," through which more current can pass when an increase in input voltage or a decrease in load current occurs, maintaining the voltage at the output nearly constant at V_Z. It should be noted that the resistor R is the element that drops the excess voltage from the input to give the desired output across the load; hence this circuit cannot compensate for voltages on the input that fall below the desired output. The selection of R is made according to the following equation:

$$R = \frac{V_{in} - V_{out}}{I_{in}} \tag{6-3}$$

where $I_{in} = I_Z + I_L$, using an average value for I_L.

6-4. DESIGN OF A VOLTAGE REGULATOR CIRCUIT

As an example, say V_{in} = 20 V, V_{out} required = 15 V over a load current range from 10 to 20 mA. Assume a 15 V Zener diode has been selected with an I_{ZT} of 6 mA. Let us determine the value of R and the rating of the diode. Assuming an average load current of 15 mA, the total current $I_{in} = I_Z + I_L = 6 + 15 = 21$ mA, as shown in Fig. 6-4.

$$R = \frac{V_{in} - V_{out}}{I_{in}} = \frac{20 - 15}{21 \text{ mA}} = \frac{5}{21 \times 10^{-3}}$$

$$= \frac{5000}{21} = 238 \text{ ohms.}$$

$$\text{Resistor power rating} = \frac{V^2}{R} = \frac{5^2}{238} = \frac{25}{238} = 0.1 \text{ W.}$$

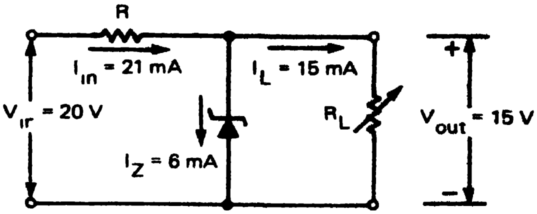
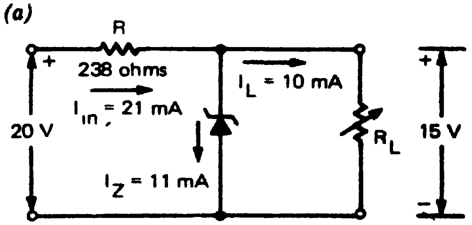
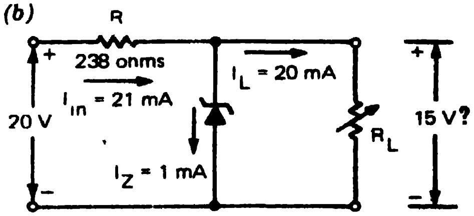


Fig. 6-4. Conditions in voltage regulator circuit at average load current.

Now let us examine the Zener diode current under minimum and maximum load current conditions, with the input voltage constant at 20 V. These conditions are depicted in Fig. 6-5.



Minimum load current



Maximum load current

Fig. 6-5. Voltage regulator circuit conditions for (a) Minimum load current, (b) Maximum load current.

Note that with this design, when the current to the load falls to 10 mA (from 15 mA, average level used) an extra 5 mA will flow through the Zener diode, giving a total of 11 mA. Most likely, the Zener voltage will still be close to 15 V. However, when the load increases to 20 mA, the extra 5 mA must be supplied from the diode, which means the Zener current must drop to 1 mA. It is doubtful the Zener voltage will still be 15 V at this low level of current and the voltage probably will be less than 15 V, since the "knee" of the curve is being approached.

Before remedying this situation, let us get some idea of the voltage variation across the diode (and load) over the current variation of 10 mA. From Fig. 6-2, the impedance of the 15 V diode at an average Zener current of 6 mA is in the order of 40 ohms. Thus, assuming the "knee" of the curve is avoided, the change in voltage that occurs is:

$$\Delta V_Z = Z_Z \Delta I_Z = 40 \times 10 \times 10^{-3} \text{ volts} = 0.4 \text{ V},$$

which corresponds to a voltage regulation of approximately 3%. Now let us see what the effect would be of selecting R so that more than the Zener test current I_{ZT} (at which V_Z is specified to be 15 V) flows through the Zener diode when the average load current is 15 mA. Use a value of 10 mA for example, instead of 6 mA. Then

$$I_{in} = I_Z + I_L = 10 + 15 = 25 \text{ mA}.$$

$$R = \frac{V_{in} - V_{out}}{I_{in}} = \frac{20 - 15}{25 \text{ mA}} = \frac{5}{25 \times 10^{-3}}$$

$$= \frac{5000}{25} = 200 \text{ ohms}.$$

$$\text{Resistor power rating} = \frac{V^2}{R} = \frac{5^2}{200} = \frac{25}{200}$$

$$= 0.125 \text{ watt}.$$

Now, when I_L increases to 20 mA, I_Z will decrease only to 5 mA, at which point the Zener voltage will still be close to 15 V. If I_L decreases to 10 mA, I_Z will increase to 15 mA. This is the worst case of power dissipation in the diode and must be checked:

$$\text{Maximum dissipation} = V_Z I_Z = 15 \text{ V} \times 15 \text{ mA}$$

$$= 225 \text{ mW}.$$

Many small glass Zener diodes have ratings of only 150 mW, so this is an important factor that must be checked.

Since the Zener diode is now operating at a higher average value of current, its dynamic impedance is in the order of 10 ohms, so that the change in voltage that will occur across the load is now only 0.1 V. This corresponds to less than 1% voltage regulation. Of course one could improve this voltage regulation even more by using a Zener diode whose value of V_Z is 15 V at an I_{ZT} of 170 mA for example, where the dynamic impedance is only 3 ohms. But this would require a 10-watt rating for the diode, with its attendant increase in cost, and would also increase the "idling" current from the supply. Alternate methods to improve voltage regulation using active devices such as transistors will be covered in a later chapter.

Notice that no matter what the value of the resistor, in each of the above cases it is used to drop 5 V to give the desired output, independent of load current, since I_{in} is constant at 21 mA in the first case and 25 mA in the second case. This is the condition where the input voltage remains constant, and only the load current varies.

6-5. EFFECT OF SUPPLY VOLTAGE VARIATIONS

Increase in Supply Voltage

The circuit of Fig. 6-4 will also regulate the load voltage if the input voltage changes. Assuming that a 400 mW Zener diode had been selected for our example (after determining that maximum dissipation would be 225 mW for the conditions specified), what maximum value of V_{in} could be tolerated without damage to the diode? Since $P = V_Z I_Z = 0.4$ W with $V_Z = 15$ V, then the maximum value of $I_Z = \frac{0.4 \text{ A}}{15} = 26$ mA. We must be sure that this Zener current value is not exceeded, so that it may only occur when the load current is at its *minimum* of 10 mA. Thus, the maximum permissible value of $I_{in} = 26 + 10 = 36$ mA. Then the maximum voltage drop across R would be

$$V_R = I_{in} R = 36 \times 10^{-3} \times 200 = 7.2 \text{ V},$$

and the maximum input voltage would be

$$V_{in} = I_{in} R + V_{out} = 7.2 + 15 = 22.2 \text{ V}.$$

The resistor power rating could be 0.5 W.

Decrease in Supply Voltage

Assuming the selected diode has an I_{ZK} of 2 mA (the Zener current near the breakover knee), what would be the minimum value of V_{in} at which the Zener diode could still maintain the output at a nominal 15 V under a 10 mA load current variation? We must now consider the Zener current at its minimum of 2 mA when the load is at its *maximum* value of 20 mA. (Why?) Under this condition, $I_{in} = 22$ mA, and the voltage drop across R is

$$V_R = I_{in} R = 22 \times 10^{-3} \times 200 = 4.4 \text{ V},$$

and the minimum permissible input voltage is

$$V_{in} = I_{in} R + V_{out} = 4.4 + 15 = 19.4 \text{ V}.$$

Thus, the supply voltage could undergo an 11% increase or a 3% decrease (from 20 V) and the Zener diode would still maintain the output voltage at a nominal 15 V. This variation could be extended by the use of a larger Zener diode, as noted earlier.

6-6. ZENER DIODE BREAKDOWN MECHANISM

Actually, the name of Zener diode is not, in general, a correct term to use, since the real Zener effect (discovered by C. E. Zener in 1934) is only found in diodes that break down below about 6 V of reverse bias. In this region of voltage, the depletion region is very narrow compared with what it is at 200 V, for example. As a result, the minority carriers do not have a large enough distance in which to gain sufficient kinetic energy to cause effective collisions. Hence, the avalanche effect is not the major cause of breakdown. However, due to the extremely narrow depletion region, in the order of only 150 – 200 Å (1 Angstrom $= 10^{-8}$ cm), there exists a high electric field, in the order of 10^6 V/cm, across the junction. This field provides a high electrical force which is responsible for "tearing" electrons out of the covalent bonds directly, rather than by collision. This is the true Zener effect. Above about 6 V, breakdown is predominantly due to avalanche multiplication as covered in Section 5-14, and diodes breaking down above that voltage should properly be called avalanche or breakdown diodes. However, the word *Zener* has been commonly used for all diodes having a sharp breakdown characteristic. The difference between these two mechanisms explains the different temperature effect on Zener diodes having a V_Z above or below 6 V, as covered in the next section.

6-7. EFFECT OF TEMPERATURE ON ZENER DIODES

We have already seen in Chapter 5 how PN junctions are temperature-sensitive in either the forward- or reverse-biased mode. Since Zener diodes operate only in the reverse mode, using the leakage current or minority carriers that are due to thermally generated electron-hole pairs, we would expect temperature to have some profound effect on the Zener voltage, V_Z. This is illustrated in Fig. 6-6 where the temperature coefficient α_Z in %/°C is shown as a function of Zener voltage for three values of Zener current.

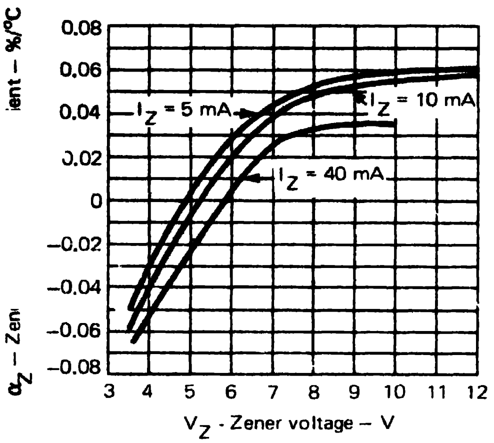


Fig. 6-6. Graph of typical Zener temperature coefficient vs. Zener voltage for silicon voltage regulator diodes. (Courtesy Texas Instruments, Incorporated.)

This curve, typical for silicon voltage regulator diodes, may be interpreted as follows. A 4 V Zener diode, for example, regardless of the level of current operation, will always be affected by temperature in a negative manner, *decreasing* in Zener voltage with an increase in temperature. On the other hand, Zener diodes with V_Z in excess of 6 V will have an *increase* in breakdown voltage with an increase in temperature. If a diode is selected with a V_Z between 5 and 6 volts and operated at the "correct" value of current, it will have very little change with temperature. The difference in sign of the temperature coefficient lies in the fact that reverse current is largely due to the true Zener effect below 6 V, but is due to the avalanche effect above 6 V. Each of these mechanisms is affected differently by temperature.

6-8. EXAMPLES OF TEMPERATURE COEFFICIENT CALCULATION

Let us determine what change in Zener voltage will take place when a 12 V and a 3.3 V diode are subjected to a change in temperature from 25°C to 110°C. Consider a 1N759, a nominal 12 V Zener diode, rated at 400 mW dissipation at 50°C, and having an $\alpha_Z = +0.060\%/°C$ at an I_{ZT} of 20 mA. Thus, at an ambient temperature (free-air temperature) $T_A = 25°C$, $V_Z = 12$ V. The change in Zener voltage for a change in temperature, ΔT, is

$$\Delta V_Z = V_Z \times \frac{\alpha_Z}{100} \times \Delta T \qquad (6\text{-}4)$$

$$= +12 \times \frac{0.060}{100} \times (110 - 25)$$

$$= +0.61 \text{ V.}$$

Thus, at $T_A = 110°C$, $V_Z = 12 + 0.61 = 12.61$ V. This is substantiated by the curves shown in Fig. 6-7. For a 1N746, a nominal 3.3 V Zener diode having an $\alpha_Z = -0.062\%/°C$ at an I_{ZT} of 20 mA, the change in Zener voltage will be:

$$\Delta V_Z = V_Z \times \frac{\alpha_Z}{100} \times \Delta T$$

$$= -3.3 \times \frac{0.062}{100} \times 85$$

$$= -0.17 \text{ V.}$$

Thus, at $T_A = 25°C$, $V_Z = 3.3$ V and at $T_A = 110°C$, $V_Z = 3.3 - 0.17 = 3.13$ V. In addition, at these elevated temperatures, the dissipation capability of the diode is reduced from 400 mW to 250 mW.

6-9. REFERENCE ZENER DIODES

The temperature coefficients for the voltage regulator diodes in the previous section are quite large compared with what can be achieved in thermally compensated Zener diodes used as reference voltage sources. In applications such as the calibration of instruments and equipment, the change in Zener voltage of the magnitude found in Section 6-8 cannot be tolerated. One method of thermal compensation is to connect a reverse-biased Zener diode with a positive temperature coefficient in series with a forward-biased diode having a negative temperature coefficient. When mounted in a single package this combination can provide a reference voltage diode having a temperature coefficient between 0.01 and 0.0005%/°C, depending upon cost, over a temperature range from -55 to +100°C. Some precision reference diodes have a stability approaching that of a standard cell. For example, an MZ605 has a V_Z of 6.2 V and a stability expressed as less than 5 ppm (parts per million)/1000 hr or as 30 μV/1000 hr. To obtain a reference or "Zener" voltage below 2 V, a diode can be biased in the forward direction, in which case it is called a stabistor, with a negative temperature coefficient. For high voltage references two or more diodes may be placed in series with the advantage of having a lower dynamic impedance than a single diode for the same voltage.

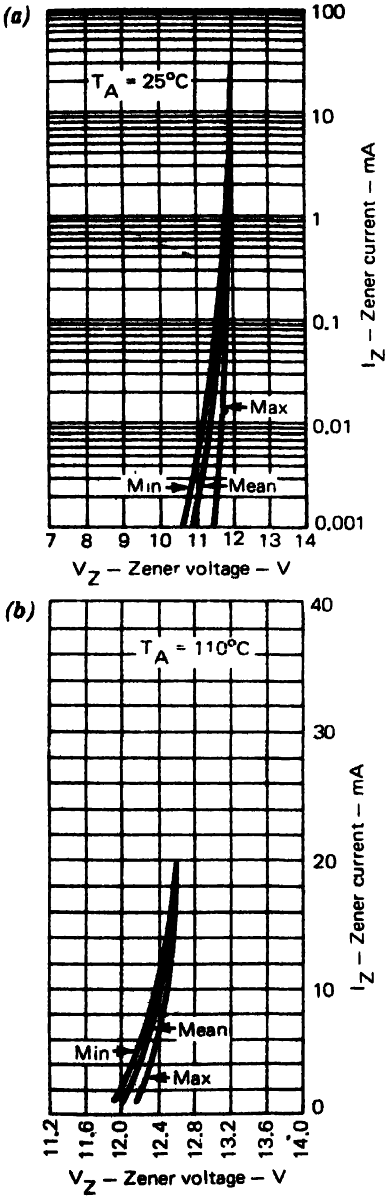


Fig. 6-7. Zener voltage vs. Zener current curves for 1N759, at ambient temperatures of (a) 25°C and (b) 110°C, showing the change in Zener voltage at an I_Z of 20 mA. (Courtesy of Texas Instruments, Incorporated.)

6-10. GAS-FILLED TUBES *

Before leaving the field of rectification, let us briefly look at the gas-filled tube. Although it is largely obsolete because of solid-state devices, some applications remain in industrial power conversion, where thousands of amperes are involved, or in some existing equipment, where characteristics similar to the Zener diode are found.

---

*This section may be omitted with no loss of continuity.

The Phanotron

The high-vacuum diode of Chapter 1 was limited in its current capability by the space charge or negative electron cloud that formed between plate and cathode. This space charge can be neutralized by adding a small amount of gas, such as mercury vapor, in the glass envelope. As the plate voltage is increased, the current through the tube increases much as in a high-vacuum diode until the ionization potential of approximately 15 volts is reached. At this point, the electrons that are being attracted to the plate have sufficient energy to collide with the mercury gas atoms and dislodge electrons. The loss of an electron leaves a positive ion, which drifts toward the cathode and tends to neutralize the electron cloud that normally builds up. This neutralization allows a much greater electron emission from the cathode, which results in increased tube current with a practically constant voltage drop across the tube of approximately 10 V.

The characteristics and symbol for the *hot-cathode* gas tube (since it has a heater) are shown in Fig. 6-8. This tube is also known as a phanotron. The applications are found where the large current demand makes semiconductor diodes not feasible, but such cases are very uncommon today, since silicon PN junction diodes are available that can handle hundreds of amperes at quite high PIV. In fact, the gas diode is limited in the PIV it can withstand due to the possibility of flashback or arc-over (reverse breakdown due to the presence of the gas in the tube) and must usually have a choke in series with the plate to limit this reverse current. Another drawback is the bombardment of the cathode by the heavy positive ions, so that a rugged structure is necessary.

The Thyratron

If a third element, called a grid, is introduced, the result is a *thyratron*, in which the grid voltage can control the "switching on" or ionizing potential of the tube, but cannot switch the tube off. However, the thyratron has largely been replaced by the solid-state silicon controlled rectifier, the SCR, which will be considered later.

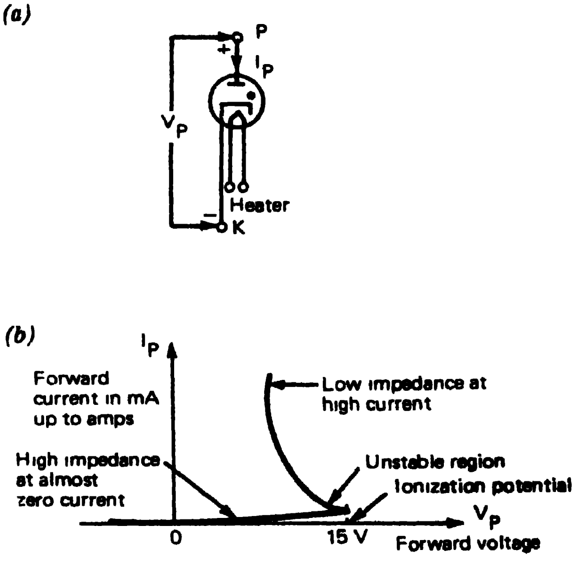

(a)
P
+
I_P
V_P
−
Heater
K
(b)
I_P
Forward current in mA up to amps
Low impedance at high current
High impedance at almost zero current
Unstable region
Ionization potential
0
15 V
V_P
Forward voltage

Fig. 6-8. (a) Symbol of a hot-cathode gas-filled tube (phanotron). (b) Characteristics of a hot-cathode gas-filled tube.

Mercury-Arc Rectifiers

High current devices, sometimes called pool-type rectifiers because of the pool of mercury that is used as a cathode, are available that have six graphite anodes, each rated at 300 A, providing a total tube capacity of 1800 A, at 150,000 V. They may be of the excitron or ignitron type, with steel cases and water-cooled jackets. Solid-state replacement seems unlikely for these devices.

Cold-Cathode Voltage-Regulator Tube

The cold-cathode gas tube is a modification of the hot-cathode gas tube, such that ionization takes place without a heater, although at a much higher plate voltage. This is made possible by using the ion bombardment of the cathode to cause secondary emission of electrons. The characteristics are very similar to the hot-cathode tube except that the ionization potential is typically 150 volts. These characteristics are shown in Fig. 6-9(a) in a slightly unusual manner.

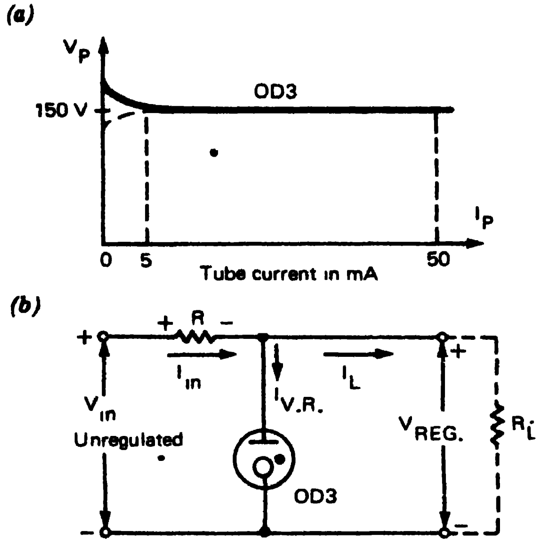


Fig. 6-9. (a) V-I characteristics of a voltage-regulator tube. (b) Circuit showing how the tube is used for voltage regulation.

Since the voltage drop across the tube is relatively constant after ionization takes place, the tube is used as a voltage regulator device, having characteristics much like a Zener diode, except the gas tube's voltage drop is in the forward direction. Provided the tube current remains between approximately 5 mA and 40 mA, the OD3 will not "go out" (lose ionization), and its rating will not be exceeded. (Rating = VI = 150 V × 40 mA = 6 watts.) The way in which the tube is used to regulate voltage against line or load variations is shown in Fig. 6-9(b). The OD3 (or OA2, etc.) acts as a "bypass valve" in the same manner as the Zener diode. The calculations pertaining to the Zener diode voltage regulating circuit also apply to the cold-cathode gas-tube diode, except that the tube current would not be allowed to drop below 5 mA. The regulation can only occur when V_{in} is above 150 volts. If V_{in} falls much below this value the OD3 is extinguished. Values of operating voltage are available for these glow-discharge diodes, as they are also known, at 75, 90, 105, and 150 volts. This range of voltage is obviously much more restricted than for the Zener diodes. For higher voltage levels, two or more diodes may be placed in series.

6-11. THE TUNNEL DIODE

If the concentration of the impurity atoms is greatly increased in a PN junction (by about 1000 times), the characteristics are completely changed, yielding a *tunnel diode*. Initially called the Esaki diode after its inventor, who developed it in 1958, this diode's characteristics are shown in Fig. 6-10(b). The explanation for the tunnel diode's volt-ampere characteristic is found in the Zener effect. Recall that a Zener diode was made to break down at a very small voltage (from 2 to 6 V) by the heavy doping causing a very narrow depletion region, in which the true Zener effect was responsible for reverse current. If the doping is increased still further, there is no reverse blocking, and the diode appears to be broken down for any reverse voltage. This increased doping also affects the forward current in the following manner.

Forward current is the sum of two components due to two different mechanisms. One, which we have already considered, arises from an external voltage reducing the potential hill or barrier across the depletion region and allowing current to flow due to majority carriers in the conduction band. This is shown

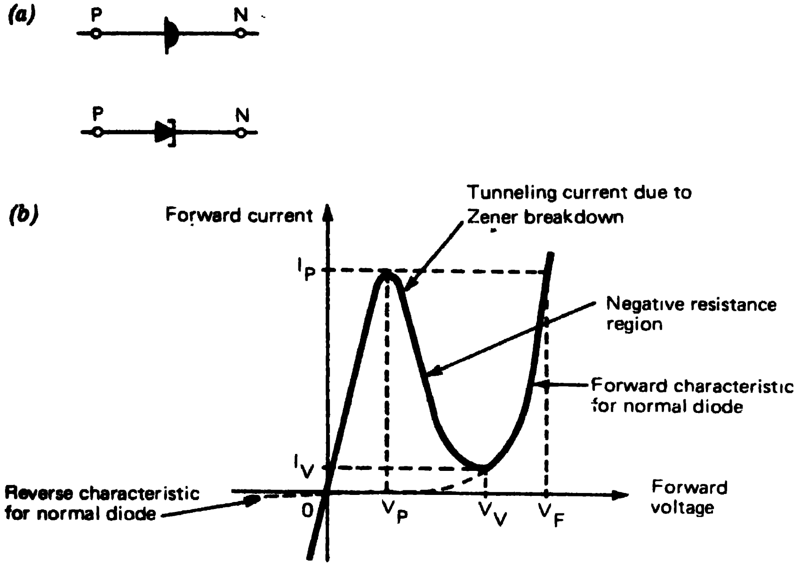

(a)
P
N
P
N
(b)
Forward current
Tunneling current due to Zener breakdown
I_P
Negative resistance region
Forward characteristic for normal diode
I_V
Reverse characteristic for normal diode
0
V_P
V_V
V_F
Forward voltage

Fig. 6-10. (a) Two commonly used symbols for a tunnel diode. (b) Tunnel diode's V-I characteristics showing the negative resistance region.

dotted in Fig. 6-10(b) and will be referred to as the normal injection current.

The other component of forward current is due to the Zener effect. The very narrow depletion region results in a very high electric field which permits carriers in the valence energy band on one side of the junction to "tunnel" through to the conduction energy band on the other side of the junction without having to overcome the potential hill. This current increases to some peak value I_P at a forward voltage V_P, which is typically 65 mV in a germanium tunnel diode and 160 mV in a gallium arsenide tunnel diode. Obviously, the normal injection current is negligible at this value of forward bias.

Because the application of forward bias reduces the electric field strength (the depletion region becomes less well defined because of the diffusion of carriers across the PN junction), the tunneling current decreases when the forward voltage exceeds V_P. This current decrease results in a region of negative conductance which is typically 8×10^{-3} mhos for a 1N3712 germanium tunnel diode. Above the valley voltage, V_v, the normal forward injection current predominates and the forward current increases in the usual manner, with $I_F = I_P$ at a forward voltage V_F of 0.5 V for a germanium tunnel diode and 0.9 V for gallium arsenide. Values of I_P range from 1 to 100 mA, with a typical I_P to I_V ratio being in the order of 6 for germanium and 10 for gallium arsenide.

6-12. APPLICATIONS OF THE TUNNEL DIODE

Initial interest in the tunnel diode was found in the area of pulse and digital circuitry. One application was as an ultra-high-speed switching device, using the tunneling mechanism, which essentially takes place at the speed of light. The response time is limited only by the diode's capacitance, which, being in the order of 1 to 10 pF, allows switching to take place (from a point at the origin to a point near the peak) with a rise time as small as 22 psec. (The rise time is the time required to go from 10% to 90% of the change in level.) The tunnel diode was also used as a logic memory storage device, taking advantage of the triple-valued feature of the curve for current between I_V and I_P. However, being a two-terminal device, it never found a very prominent place in digital computer circuitry compared with the greater flexibility (and comparable speed) of the three-terminal transistors now available.

A small-signal equivalent circuit for the tunnel diode is shown in Fig. 6-11.

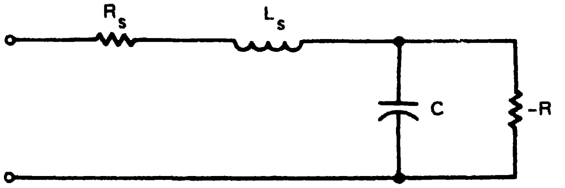


Fig. 6-11. Small-signal equivalent circuit of a tunnel diode.

R_s is typically 1 to 5 ohms, L_s from 0.1 to 4 nH, and C from 0.35 to 100 pF. The extremely low capacitance and inductance permit its use in microwave oscillators at frequencies in the order of 10 GHz. The negative resistance of the tunnel diode allows its use in relaxation oscillators, and in this respect it is very similar to the unijunction transistor, which will be covered in a later chapter.

6-13. THERMISTORS

Although a thermistor is not a rectifying device, it is a semiconductor and may provide a voltage-regulating function for ac voltages that is not possible with a Zener diode. We saw the effect of an increase in level of doping in a semiconductor, yielding the Zener and tunnel diodes. Now consider intrinsic germanium (no impurity). This material has such a very large negative temperature coefficient that a 3° C change causes a resistive change of about -13%. This very large change is used to advantage in "thermal resistors" or thermistors, practical versions of which consist of mixtures of metallic oxide, such as manganese, nickel, cobalt, copper, iron and uranium. They are manufactured in the form of beads ranging in diameter from 0.006 inch to 0.1 inch, probes (beads sealed in glass rods, discs or washers) up to 1 inch in diameter, and rods from 1/4 to 2 inches long. Resistance values of thermistors are available from 100 ohms to over 10 megohms at room temperature, depending upon type. Fig. 6-12(a) shows how the resistance varies with temperature for a typical thermistor. It is this very large change in resistance which makes the thermistor very valuable as a temperature measurement (and control) device, some circuits being capable of controlling temperature to within ±0.0005°C.

Temperature Coefficient

The change in resistance with temperature is expressed by the temperature coefficient, α, defined as follows:

$$\alpha = \frac{1}{R_o} \frac{\Delta R_o}{\Delta T} \text{ ohms/ohm/}^\circ\text{C}, \qquad (6\text{-}5)$$

where R_o is the resistance with zero power in the thermistor at 25°C. Alternatively, α may be in %/°C, and is typically -4%/°C at room temperature, dropping to -2%/°C at 200°C.

Time Constant

The time constant is the time required for a thermistor to change its temperature 63% of the way from its original value to the new ambient temperature impressed upon it in a step change. Values range from 1 second for small bead types to 300 seconds for some large discs. This is a serious limitation in some applications, but is used to advantage in others such as current surge suppression or time delays extending from a fraction of a second to several hours.

V-I Static Characteristics

The way in which voltage and current are related for a given thermistor is usually shown by a curve drawn on logarithmic scales, as shown in Fig. 6-12(b).

(a)

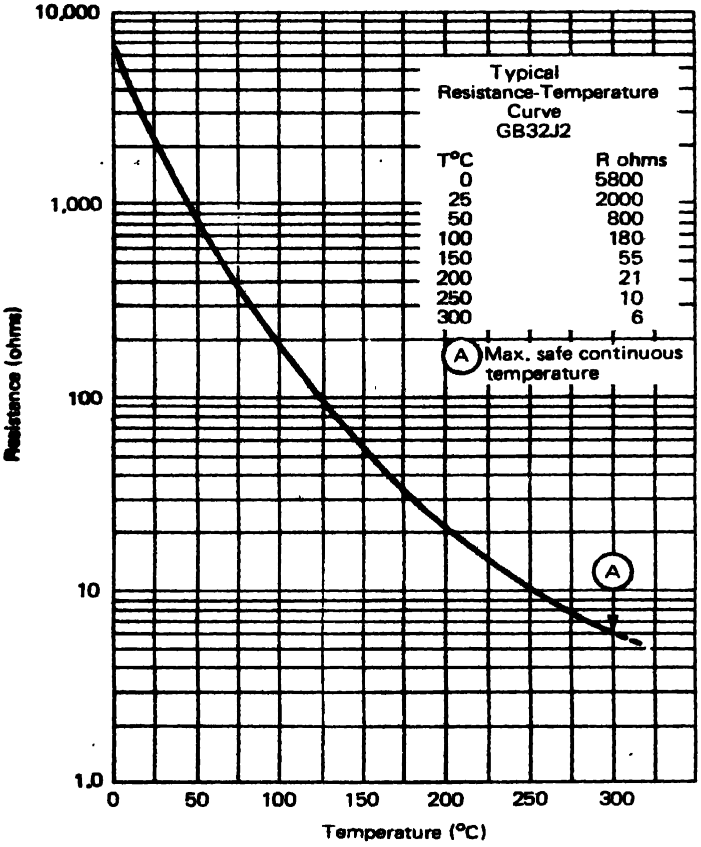


Fig. 6-12. (a) Typical resistance-temperature curve for a thermistor.

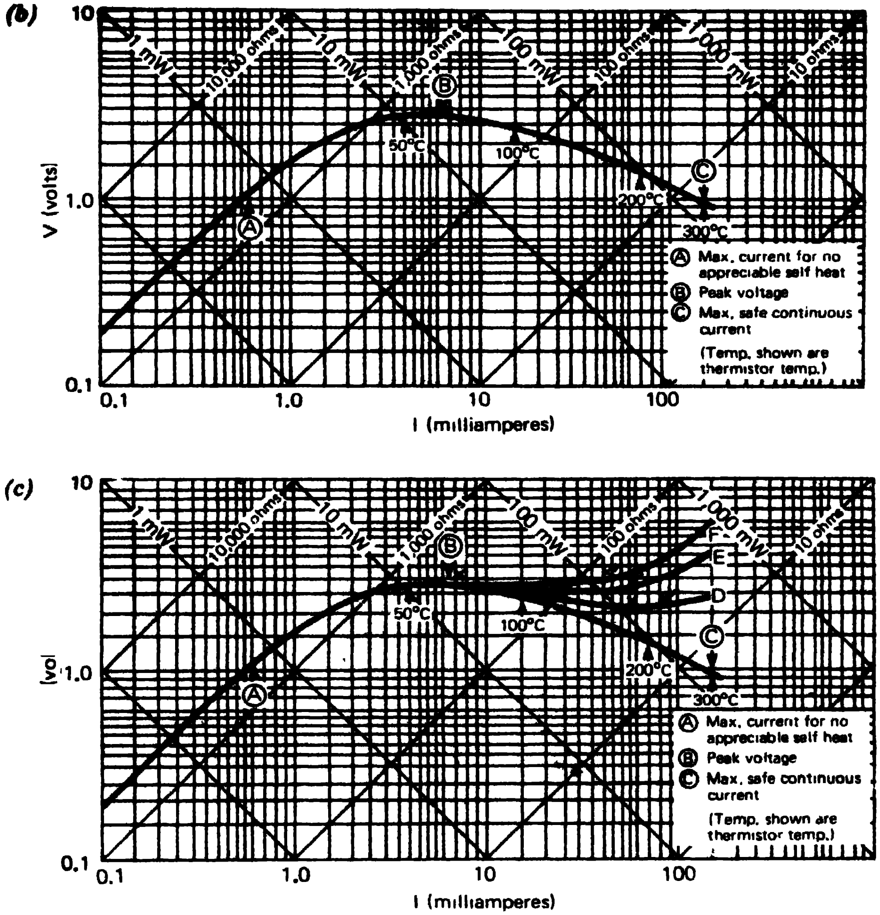


Fig. 6-12
(cont'd)
(b) Typical V-I curve obtained for the thermistor shown in (a) at an ambient temperature of 25°C.
(c) V-I curves drawn to show the effect of adding a series resistor to the thermistor in (b). Curves D, E, and F are for series resistors of 10 ohms, 20 ohms and 30 ohms respectively.
(Courtesy Fenwal Electronics Division of Walter Kidde & Company, Inc.)

There are two distinct regions in this curve with different applications, as described in the following two sections.

6-14. EXTERNALLY HEATED THERMISTOR OPERATION

The portion of the curve up to point A indicates that if the power dissipated in the thermistor is less than 1 mW then V and I obey Ohm's Law. That is, the resistance is constant (at approximately 2 K ohms). In this region, the current through the thermistor is not sufficient to raise the temperature of the device appreciably above ambient, and the variation in resistance will depend only upon changes in the temperature of the surroundings as indicated in Fig. 6-12(a). It is this linear region that is used for temperature measurement, in which the thermistor may be placed in one arm of a Wheatstone bridge, and an unbalance due to temperature (and resistance) change will be indicated on a suitably calibrated ammeter or voltmeter. (This application is explored further in Chapter 2S.) The resistance change with temperature could be shown on the V–I characteristic of Fig. 6-12(b) and would consist of almost parallel curves displaced from the curve for 25°C by an amount dependent upon the change in temperature. The curves would move upward and to the left for a drop in temperature and downward to the right for an increase in temperature, as indicated by the diagonal resistance lines.

6-15. SELF-HEATED THERMISTOR OPERATION

Referring to the V–I characteristics of Fig. 6-12(b), it is seen that an increase in current past point A causes a reduction in thermistor resistance (the curve approaches more closely the 1000-ohm line). This resistance decrease is due to the heat developed in the thermistor itself by the action of current through it. A peak value of voltage occurs at B (at which the thermistor body temperature is above 50°C), and increasing the current causes so much self-heating that the resulting drop in resistance causes a drop in thermistor voltage. The region from B to C exhibits a negative resistance effect, and current will continue to increase until a steady state is reached where the thermistor dissipates as much power as is supplied to it. A series resistor is often used to prevent the current from becoming excessive and to limit the maximum thermistor temperature to 300°C for the type shown in Fig. 6-12(b).

6-16. APPLICATIONS IN THE SELF-HEATED MODE

The applications in the self-heated mode are even more numerous and diverse than the measurement and control of temperature in Section 6-14. Under conditions of self-heating, the thermistor is very sensitive to the *rate of heat removal* and so may be used as a flow measurement device capable of measuring flow rates as low as 0.001 cubic cm per minute. A circuit used for this application is shown in Fig. 6-13(a). Here one thermistor is inserted in the line of fluid flow and the other thermistor is used to compensate for temperature variations. That is, we want bridge unbalance to result from the variation in cooling depending upon fluid flow and not from fluid or ambient temperature variations. Such a flow measurement device, utilizing a thermistor probe in the fuel line of an automobile could be used as an "instantaneous" fuel consumption indicator, or, if combined with a voltage proportional to speed could give an "instantaneous" miles-per-gallon measurement.

Alternatively, a vacuum gauge would result if the sensing thermistor were placed in a vacuum, since the lack of air would reduce cooling and cause bridge unbalance, and the meter could be calibrated directly in mm of mercury. A gas analysis instrument for gas chromatography could be obtained if the reference thermistor were surrounded by a reference gas, such as air for example. If the gas flowing over the sensing thermistor were pure air, then the bridge would be balanced. If the gas had a content of carbon dioxide, which has a lower thermal conductivity than air, the sensing thermistor would become hotter, and the amount of unbalance of the bridge would be proportional to the carbon dioxide content. It is possible to make such an instrument to have a full scale reading of 1/2% CO_2 in air.

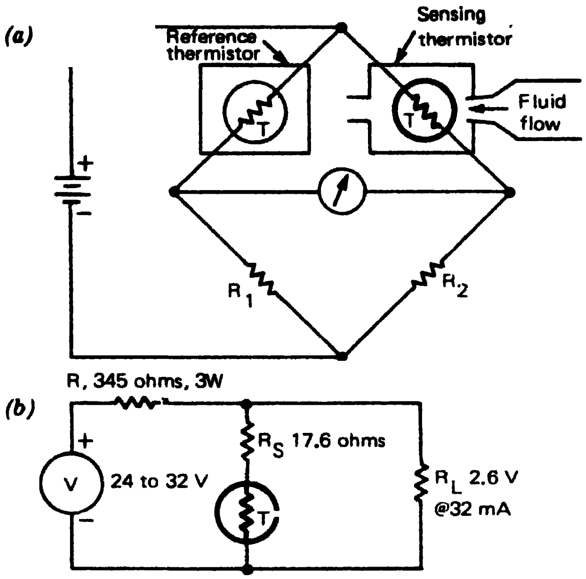


Fig. 6-13. (a) Wheatstone bridge circuit for flow measurement, vacuum gauge, or gas analysis instrument. (b) Voltage regulating circuit for dc or ac voltage.

6-17. AC VOLTAGE REGULATOR

An interesting application in the self-heated mode is found in a voltage regulating circuit that can regulate an *ac voltage* in the face of line or load variations. This is possible because the thermistor is not polarity sensitive as is a Zener diode, which can only directly regulate dc voltages. Such a circuit is shown in Fig. 6-13(b), where a resistor in series with the thermistor acts as the controlling element. The way in which regulation is achieved is illustrated by the curves drawn in Fig. 6-12(c). The effect of adding a resistor in series with the thermistor is to create a flat portion, X – X′ where voltage across the combination is relatively insensitive to current variations through the resistor and thermistor. The particular value of resistance to use depends upon the value of load voltage desired and the range of current operation. By suitable selection of thermistor and resistor, any voltage between 1/2 volt and 100 volts may be regulated, with voltage regulation in the order of 1%. In the circuit of Fig. 6-13(b), the line voltage varies between 24 and 32 volts, but the load voltage is "constant" at 2.6 volts ±0.02 volt, at a load current of 32 mA. This represents a voltage regulation of ±0.77%. With the thermistor used above, having a 2-second time constant, the response to a change in voltage is quite quick. For some larger thermistors having a greater current capability, the large time constant of several minutes may be a disadvantage although such thermistors have quite a flat response region even without any series resistance.

A circuit that may be used as a combined voltage regulator and temperature controller is shown in Fig. 6-14.

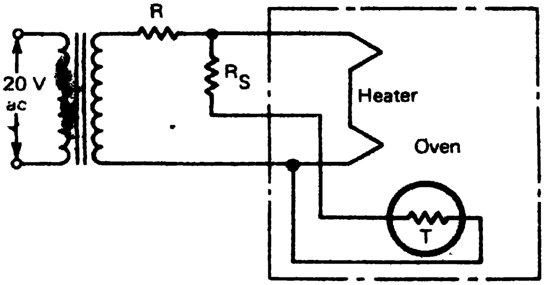


Fig. 6-14. Combined voltage-regulator and temperature-controller circuit.

Here the thermistor is located in the oven which is to be held at a given temperature by the heater. If the line voltage increases, tending to increase the heater voltage and oven temperature, more current flows through the thermistor, with a resulting decrease in resistance. Less current will flow through the heater, and the temperature will tend to hold steady. If for any reason there is a drop in temperature of the oven, due to a drop in ambient temperature for example, the thermistor will increase in resistance, causing greater heater current to compensate for the drop. Temperature-controlled ovens are often used to control the temperature of a crystal, whose frequency of oscillation may be very dependent upon temperature. Many crystals need very close temperature control to maintain the proper operating frequency of transmitters in communication equipment.

6-18. SUMMARY

(1) A Zener diode is a PN junction operated in the reverse-biased mode to take advantage of its sharply defined breakdown voltage.

(2) The Zener voltage V_Z is specified at some test value of current I_{ZT}, at which the diode will exhibit some dynamic impedance $Z_Z = \frac{\Delta V_Z}{\Delta I_Z}$, which depends upon the Zener voltage of the diode and the level of Zener current.

(3) A Zener diode may be used to regulate the load voltage at the value V_Z by acting as a "bypass valve" to counteract line voltage or load current variations.

(4) Diodes having a breakdown voltage below about 6 V rely on the true Zener effect (high electric field moves electrons from bonds), while the avalanche effect is responsible for reverse current above 6 V.

(5) Zener diodes have a temperature coefficient, α_Z, which generally is negative for V_Z below about 6 V but positive above 6 V, and is expressed in percent of V_Z per °C, with the change in Zener voltage given by the equation:

$$\Delta V_Z = V_Z \times \frac{\alpha_Z}{100} \times \Delta T.$$

(6) Reference Zener diodes are available with α_Z as low as 0.0005%/°C.

(7) The admission of a small amount of mercury gas increases the current capability of a hot-cathode gas-filled tube.

(8) A cold-cathode or glow-discharge diode may be used as a dc voltage regulator in a similar manner to a Zener diode.

(9) A tunnel diode's characteristics exhibit a negative resistance region and it may be used as an ultra-fast switching device or in ultra-high-frequency oscillators.

(10) Thermistors are semiconductors which have a high negative temperature coefficient in the order of −4%/°C and may be used for temperature measurement and control.

(11) Thermistors may also be used in the self-heated mode where they are sensitive to the *rate* of heat removal and have applications such as flow measurement devices or voltage regulators for dc or ac.

QUESTIONS

Q6-1. What is the polarity of voltage across a Zener diode that is being used in a voltage regulating circuit? Draw the symbol.

Q6-2. What range of Zener voltages are available in Zener diodes? How may one obtain a higher Zener voltage? What advantage would this have over the use of a single Zener diode?

Q6-3. Explain, with reference to a Zener diode's characteristic curve, what you understand by the following: I_{ZK}, I_{ZT}, Z_Z.

Q6-4. How does the dynamic impedance of a Zener diode vary with (a) Zener current? (b) Zener voltage?

Q6-5. What is the significance of the dynamic impedance of a Zener diode? What would the impedance be for an ideal Zener diode? What result would this have?

Q6-6. Referring to Fig. 6-3, describe the voltage regulating capability of the circuit for the following conditions:

(a) The load resistance R_L increases.
(b) The input voltage decreases.

Q6-7. For what conditions does the Zener diode in Fig. 6-3 have maximum dissipation?

Q6-8. What factors determine (a) The maximum decrease in supply voltage, V_{in}, in Fig. 6-3 before effective voltage regulation is lost?

(b) The maximum permissible increase in V_{in}?

Q6-9. Explain briefly the two different mechanisms at work in Zener diodes below and above approximately 6 V breakdown.

Q6-10. What approximate values of Zener voltage and current provide a minimum temperature effect on the Zener voltage?

Q6-11. What is a stabistor?

Q6-12. A cold-cathode voltage-regulator tube can provide a very similar regulating function to that of the Zener diode. What are some of the major differences and disadvantages of the cold-cathode tube compared with the Zener diode?

Q6-13. (a) Explain how a tunnel diode may be thought of as a heavily doped Zener diode.

(b) Does a tunnel diode have a rectifying property?

(c) Give two applications of a tunnel diode.

Q6-14. What basic semiconductor property does a thermistor rely upon?

Q6-15. What is meant by the time constant of a thermistor?

Q6-16. (a) When does a thermistor obey Ohm's Law?

(b) What applications does the thermistor have in this region?

Q6-17. (a) What is the thermistor sensitive to when operated in the self-heated mode?

(b) What are some of the applications in this region of operation?

Q6-18. (a) What permits a thermistor to provide voltage regulation for ac in contrast with a Zener diode?

(b) What disadvantage does a thermistor have when used in a voltage regulating circuit?

PROBLEMS

P6-1. (a) Use the curves in Fig. 6-2 to determine the dynamic impedance of a 10 V Zener diode operating at a current of 5 mA.

(b) Determine the change in voltage across the Zener diode in (a) if the current through the diode changes by 2 mA.

(c) If the change in voltage in (b) is higher than can be tolerated, what must be the new diode current for only 10 mV change to occur when the current changes by 2 mA?

P6-2. Design a voltage regulator circuit to provide 12 V across a load whose current varies from 5 mA to 35 mA. An unregulated 18 V dc source is to be used. Assume a Zener diode has been chosen with a Zener voltage of 12 V @ $I_{ZT} = 20$ mA. Specify:

(a) The voltage dropping resistor and its power rating using an average value for I_L when $I_Z = I_{ZT}$.

(b) The minimum power rating of the Zener diode.

(c) The maximum value of I_{ZK} to maintain reasonable voltage regulation when load current is at its maximum.

(d) The approximate voltage variation and percent regulation expected from 5 mA to 35 mA, with the aid of Fig. 6-2 and the average Zener current.

P6-3. For the voltage regulator circuit designed in P6-2, determine:

(a) The maximum increase in supply voltage that can be tolerated if the diode is rated at 1 W.

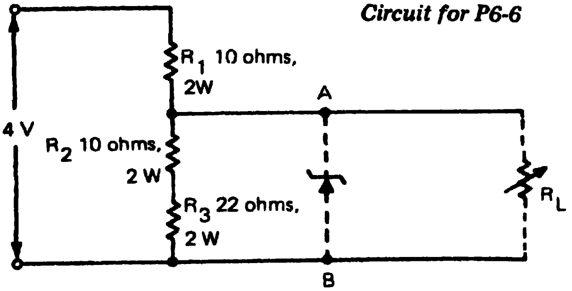
(b) The minimum input voltage if the diode has an I_{ZK} of 2 mA.

P6-4. A 10 V Zener diode is operated at a reverse current of 5 mA at room temperature of 20°C.

(a) Determine, with the use of Fig. 6-6, what its Zener voltage will be at a temperature of 75°C.
(b) Assuming the change in Zener voltage in (a) is more than can be tolerated, what change will take place at 75°C if the diode is operated at 40 mA?
(c) What would be the overall change in Zener voltage at 75°C if two 5 V Zener diodes were operated in series at a current of 5 mA?
(d) What would be one disadvantage of the operation in (c) compared with that in (b)?

P6-5. In what range of current should a 4 V Zener diode be operated so that a 1 mA change in current causes a maximum change of 40 mV across the diode *and* the Zener voltage does not fall below 3.9 V when the temperature rises 50°C?

P6-6. A student decides to build a voltage divider so that he can run his 9 V tape recorder off his car voltage system. He selects resistors as shown below:



Circuit for P6-6

(a) Without any Zener diode or load R_L connected, what is the voltage from A to B with the car's generator (or alternator) supplying 14 V?

(b) Deciding to apply his new-found knowledge, the student decides a voltage regulator employing a 9.1 V Zener diode would be a good idea. How much current flows through the Zener diode when connected from A to B and what should be its power rating? (Assume no load connected yet.)
(c) What is the maximum load, R_L, that can be applied without losing effective voltage regulation if the current through the Zener diode must be at least 25 mA to maintain a 9.1 volt drop across it?
(d) If the temperature in the car rises from 70°F to 115°F, what will be the voltage across the tape recorder while it is playing at full volume (maximum load in (c)).

P6-7. Use the curve in Fig. 6-12(a) to determine the average temperature coefficient of the GB32J2 thermistor over the range of temperature from (a) 25°C to 50°C, (b) 200°C to 250°C.

P6-8. A thermistor at 25°C and with negligible current through it has a resistance of 1 M ohm. What will be the change in resistance and the new resistance at 50°C assuming the temperature coefficient is constant at -3%/°C?

P6-9. A 2-K ohm thermistor at 25°C has a time constant of 45 seconds.

(a) What will be its temperature 45 seconds after being placed in a 125°C furnace?
(b) Approximately how long will it take to cool back to room temperature at 25°C after it had been in the furnace for 30 minutes?

P6-10. An ac voltage regulating circuit is to be built similar to that in Fig. 6-13(b). The thermistor is a GB32J2, and R_s is a 30-ohm resistor, to provide the F curve shown in Fig. 6-12(c). The load voltage is to be maintained at 3 V over a load current variation from 5 mA to 35 mA. The input is 6.3 V.

(a) Select R so that the thermistor-resistor current is 10 mA when the load current is maximum. Specify the power rating of R. (Note that the logarithmic scale for V progresses in 0.5 V intervals up to 6 V, then in 1 V intervals; similarly, I progresses from 10 mA in 5 mA intervals up to 60 mA, then in 10 mA intervals.)
(b) What is the total resistance of the thermistor-resistor combination at maximum load current? What is the thermistor's resistance? What is the voltage drop across the thermistor, and what is its power dissipation?
(c) What is the output voltage when the load current drops to minimum? What percent voltage regulation does this represent?
(d) Repeat part (b) with the load current at minimum.
(e) Does the voltage drop across the thermistor at minimum load current match the expected voltage from the use of the curve in Fig. 6-12(b)? Estimate the temperature of the thermistor.
(f) If the input voltage should increase to 7.5 V, what will be the load voltage at minimum load current of 5 mA? (Requires a trial and error solution.)
(g) Determine the minimum voltage that the input can drop to under maximum load conditions and still maintain the output at 3 V.

PART 2

GENERAL AMPLIFIER CHARACTERISTICS

7

GENERAL AMPLIFIER CHARACTERISTICS I: DISTORTION

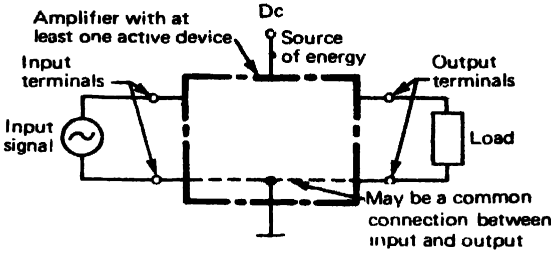

Amplifier with at least one active device
Dc
Source of energy
Input terminals
Output terminals
Input signal
Load
May be a common connection between input and output

Fig. 7-1. Black-box representation of an amplifier.

7-0. INTRODUCTION

The purpose of this chapter is to introduce some of the general characteristics that pertain to amplifiers of the tube or solid-state type. The general concept of amplification is considered along with current, voltage, and power gains. Since an amplifier generally produces an enlarged version of an input signal, the fidelity with which this is done is important, so that a general discussion of distortion and its calculation and measurement is included.

7-1. CONCEPT OF AMPLIFICATION

The amplifier may be thought of as a "black box" that has two input terminals for the signal that is to be amplified, two output terminals for the connection of the load, and a means of supplying power to the amplifier. One of the input and output terminals may be common—a straight through connection—, but not always. (This connection may be tied to the amplifier chassis and/or to the power line ground.) See Fig. 7-1. The signal to be amplified may be an ac or dc voltage, but, since many dc amplifiers often convert the dc to a proportional ac voltage (and then reconvert it to dc after amplification), we will assume in this discussion that the input and output are alternating in nature, at some fixed or variable frequency.

By input signal we mean a low level voltage such as that obtained from a phonograph cartridge, tape head, or a transducer such as a thermocouple, pressure gauge, etc. The output load may be a loudspeaker in an audio amplifier, a motor in a servo amplifier, a relay in a control application, etc. In any event, the output of the amplifier is an enlarged version of the input. To amplify means to increase the size of, raise the level of, or magnify, the input. In order to do this, all amplifiers need a source of energy. This may be provided by a battery (usually a dry battery) or a dc source resulting from a rectifier and filter combination as covered in Chapters 2, 3 and 4. In this case, the amplifier is not truly portable and usually operates from a 120 V, 60 Hz outlet.

The amplifier also contains at least one active device. This may be an electron tube or a semiconductor device (usually with three or more terminals) that can provide a control function. Basically, the active device converts the energy from the dc source into energy at the output of the amplifier that is proportional to the input signal. The ac input signal merely supplies a means of controlling the dc to ac conversion that takes place in the tube or transistor and the control is usually accomplished with comparatively little input signal power.

7-2. AMPLIFIER NOTATIONS

In order to refer to the amplification that takes place, let us use the notations shown in Fig. 7-2.

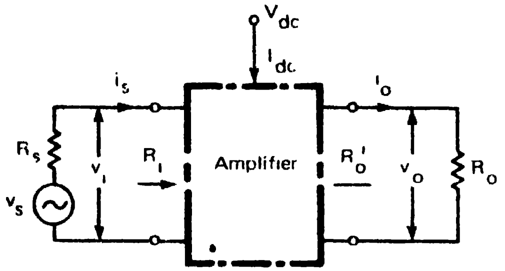


Fig. 7-2. Circuit showing amplifier current, voltage, and resistance notations.

v_s is the open circuit source or signal voltage = $V_{sm} \sin \omega t = \sqrt{2}V_s \sin \omega t$

i_s is the source or signal current = $I_{sm} \sin \omega t = \sqrt{2}I_s \sin \omega t$

R_s is the internal resistance of the source, ohms

R_i is the amplifier input resistance, ohms

v_i is the amplifier input voltage = $V_{im} \sin \omega t = \sqrt{2}V_i \sin \omega t$

v_o is the amplifier output voltage = $V_{om} \sin \omega t = \sqrt{2}V_o \sin \omega t$

i_o is the amplifier output current = $I_{om} \sin \omega t = \sqrt{2}I_o \sin \omega t$

R_o is the load resistance, ohms

R_o' is the amplifier output (or internal) resistance, ohms

V_{dc} is the dc supply voltage

I_{dc} is the dc supply current

7-3. CURRENT GAIN, A_i

The ratio of output current to input current is the current gain:

$$A_i = \frac{i_o}{i_s} = \frac{I_o}{I_s}. \tag{7-1}$$

It may also be obtained by using the ratio of the peak-to-peak values:

$$A_i = \frac{I_{o,p\text{-}p}}{I_{s,p\text{-}p}}. \tag{7-2}$$

7-4. VOLTAGE GAIN, A_v

The ratio of output voltage to input voltage is the voltage gain:

$$A_v = \frac{v_o}{v_i} = \frac{V_o}{V_i} = \frac{V_{o,p\text{-}p}}{V_{i,p\text{-}p}}. \tag{7-3}$$

Sometimes the overall loaded voltage gain may be required; it would be the ratio between V_o and V_s.

7-5. POWER GAIN, A_p

The ratio of power delivered to the load to power delivered to the amplifier is the power gain:

$$A_p = \frac{P_o}{P_i} \tag{7-4}$$

$$= \frac{V_o I_o}{V_i I_s} \tag{7-5}$$

$$= A_v \cdot A_i \tag{7-6}$$

and is clearly the product of the voltage and current gains.

An amplifier may or may not exhibit both a voltage and a current gain, but in general will show a power gain. However, whether voltage or power gain is the more important depends upon the application. Power gain may also be expressed in decibels, as explained in Section 8-1.

Example 7-1. An amplifier has a signal input voltage, V_i, of 0.25 V and draws 1 mA from the source. The amplifier delivers 8 V to a load at 10 mA. Determine the current, voltage and power gains.

Solution

$$A_i = \frac{I_o}{I_s} = \frac{10\text{ mA}}{1\text{ mA}} = 10$$

$$A_v = \frac{V_o}{V_i} = \frac{8\text{ V}}{0.25\text{ V}} = 32$$

$$A_p = A_v \cdot A_i = 32 \times 10 = 320.$$

Alternatively,

$$P_i = V_i I_s = 0.25 \times 1 \times 10^{-3} = 0.25\text{ mW}$$

$$P_o = V_o I_o = 8 \times 10 \times 10^{-3} = 80\text{ mW}$$

$$\therefore A_p = \frac{P_o}{P_i} = \frac{80\text{ mW}}{0.25\text{ mW}} = 320.$$

Note that the gains are pure numbers: they have no units, since they are ratios of the same dimension.

7-6. AMPLIFIER INPUT RESISTANCE, R_i

The amplifier presents a load, R_i, to the source which may be a very high value of resistance in tube and field-effect transistor amplifiers, but may range from only 50 to a few thousand ohms in bipolar transistor amplifiers. The effect of R_i on overall voltage gain may be seen in Fig. 7-3.

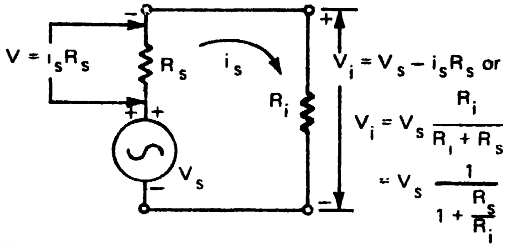

$V = i_s R_s$
R_s
i_s
R_i
V_s
+
−
$V_i = V_s - i_s R_s$ or
$V_i = V_s \frac{R_i}{R_i + R_s}$
$= V_s \frac{1}{1 + \frac{R_s}{R_i}}$

Fig. 7-3. Circuit showing how amplifier input resistance reduces input voltage.

The input voltage to the amplifier will be reduced from V_s by an amount that is dependent upon both R_s and R_i. In fact the reduction depends upon their ratio, $\frac{R_s}{R_i}$, as shown by using the voltage divider relationship:

$$V_i = V_s \times \frac{R_i}{R_i + R_s} \qquad (7\text{-}7)$$

$$= V_s \times \frac{1}{1 + \frac{R_s}{R_i}} \qquad (7\text{-}8)$$

Thus, if $R_i >> R_s$ (negligible loading), the input voltage to the amplifier is substantially the same as the open circuit source voltage.

If R_s is comparable to R_i (10% of R_i or more), the smaller voltage at the input to the amplifier will result (for a given amplifier voltage gain) in a smaller output voltage than if the input resistance had been infinite (or if R_s had been zero).

The above equation should indicate an experimental method of obtaining the input resistance of an amplifier if the internal resistance of the source is known, by noting the open circuit voltage V_s and then the voltage V_i after connection to the amplifier. Another method of determining R_i in the laboratory is to insert a resistor of known value between the source and the input to the amplifier and measure the voltage drop across it. This will yield the source current, I_s, from which the input resistance $R_i = \dfrac{V_i}{I_s}$.

Example 7-2. Determine the input resistance of the amplifier in Example 7-1 when an input voltage V_i of 0.25 V draws 1 mA from the source.

Solution

$$R_i = \frac{V_i}{I_s} = \frac{0.25\ \text{V}}{1 \times 10^{-3}\,\text{A}} = 250\ \text{ohms}.$$

Example 7-3. What must be the open circuit voltage of the source, V_s, in Example 7-2 to provide an amplifier input voltage V_i of 0.25 V when the internal resistance of the source is 50 ohms? Refer to Fig. 7-4.

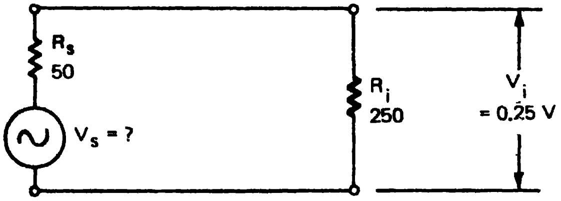


Fig. 7-4. Circuit for Example 7-3, showing the source and input resistance.

Solution

$$V_i = V_s \frac{1}{1 + \dfrac{R_s}{R_i}}$$

$$\therefore V_s = V_i \left(1 + \frac{R_s}{R_i}\right)$$

$$= 0.25 \left(1 + \frac{50}{250}\right)$$

$$= 0.25 \times 1.2$$

$$= 0.3\ \text{V}.$$

Thus the open circuit source voltage must be 0.3 V.

Example 7-4. If the amplifier's input resistance is now increased (by modifying the amplifier) to 5 K ohms, what will be the new output voltage from the amplifier (assuming its voltage gain remains at 32, the value calculated in Example 7-1), with V_s = 0.3 V and R_s = 50 ohms?

Solution

$$V_i = V_s \times \frac{1}{1 + \dfrac{R_s}{R_i}} = 0.3 \times \frac{1}{1 + \dfrac{50}{5000}}$$

$$= 0.3 \times \frac{1}{1.01} = 0.297\ \text{V}$$

$$\therefore V_o = AV_i = 32 \times 0.297 = 9.5\ \text{V}.$$

This new output of 9.5 V with an input resistance of 5 K is to be compared with an output of 8 V with an input resistance of 250 ohms; the open circuit source voltage remained the same at 0.3 V and R_s at 50 ohms.

7-7. AMPLIFIER OUTPUT RESISTANCE, R_o'

This is essentially the internal resistance of the amplifier, such that an equivalent circuit for the amplifier output with a load connected would be as shown in Fig. 7-5.

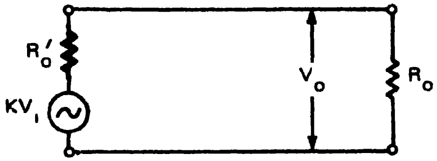


Fig. 7-5. Equivalent circuit for the output of an amplifier.

The open circuit voltage of the amplifier output would be KV_i, where K is a constant for a given amplifier. Connection of a load R_o would provide an output voltage V_o, somewhat smaller than KV_i, depending upon the relative values of R_o' and R_o as shown by the following:

$$V_o = KV_i \times \frac{R_o}{R_o + R_o'} \tag{7-9}$$

$$= KV_i \times \frac{1}{1 + \frac{R_o'}{R_o}} . \tag{7-10}$$

The variation in gain of the amplifier with R_o is shown by the following equation:

$$A_v = \frac{V_o}{V_i} \tag{7-11}$$

$$= \frac{K}{1 + \frac{R_o'}{R_o}} . \tag{7-12}$$

The variation in voltage gain with R_o for a given value of R_o' may be plotted on semilogarithmic paper which accommodates a large span of values for R_o horizontally but has a linear scale vertically for A_v. Fig. 7-6 shows a plot of A_v vs. $\frac{R_o}{R_o'}$.

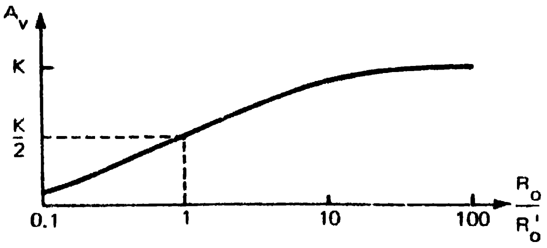


Fig. 7-6. Graph of amplifier voltage gain as a function of load resistance to output resistance ratio.

It is to be noted that when the amplifier is open-circuited ($R_o = \infty$), A K, and when $R_o = R_o'$, $\left(\frac{R_o}{R_o'} = 1\right)$, $A_v =$ This affords a method of determining an amplifier's output resistance.

Ideally, R_o' should be zero, so that the output voltage is independent of the load resistance. This is in contrast to the input resistance, which ideally would be infinite.

7-8. MAXIMUM POWER TRANSFER

If an amplifier is being used to provide as large a voltage at the output as possible, then the load R_o should be made as large as is practical compared with R_o'. This is shown in the previous section. But if maximum power transfer is desired as in power amplifiers, then the load R_o should be matched to R_o'. Consequently, we often find matching transformers on the output of an amplifier for this purpose. For the same reason, transformers are often used to couple one stage of amplification to the next, where, in transistor amplifiers especially, relatively small input resistances are not uncommon. Examples of these considerations are covered in the following chapters.

7-9. CONVERSION EFFICIENCY

It has been stated that an amplifier is essentially an energy conversion device, with the input signal providing the control but the dc source providing the major portion of the energy. The conversion efficiency for the active device in a single stage of amplification is defined as follows:

$$\eta = \frac{\text{ac signal power delivered to the total load}}{\text{dc input power to the active device}}$$

$$= \frac{P_o}{P_{dc}} . \tag{7-13}$$

This may be referred to as the collector circuit efficiency in a transistor amplifier or as the plate circuit efficiency in a tube amplifier. Maximum theoretical values for conversion efficiency depend upon the way in which the load is coupled to the active device and the class of operation of the amplifier, and may be anywhere from 25% to 90%, as developed in later chapters.

Example 7-5. A single-stage amplifier employing one active device is powered by a 9 V battery which has a current drain of 20 mA. If the load voltage is 3 V at 12 mA, determine the conversion efficiency of the active device.

Solution

$$P = VI = 3 \cdot 12 \times 10^{-3}\ \mathrm{W} = 36\ \mathrm{mW}$$

$$P_{dc} = V_{dc} I_{dc} = 9 \cdot 20 \times 10^{-3}\ \mathrm{W} = 180\ \mathrm{mW}$$

$$\eta = \frac{36}{180} \times 100\% = 20\%$$

7-10. CLASSES OF AMPLIFIER OPERATION

The class of operation of an amplifier is determined by the period of conduction of current through the active device compared with the input signal's period. There are three main classes as shown in Fig. 7-7.

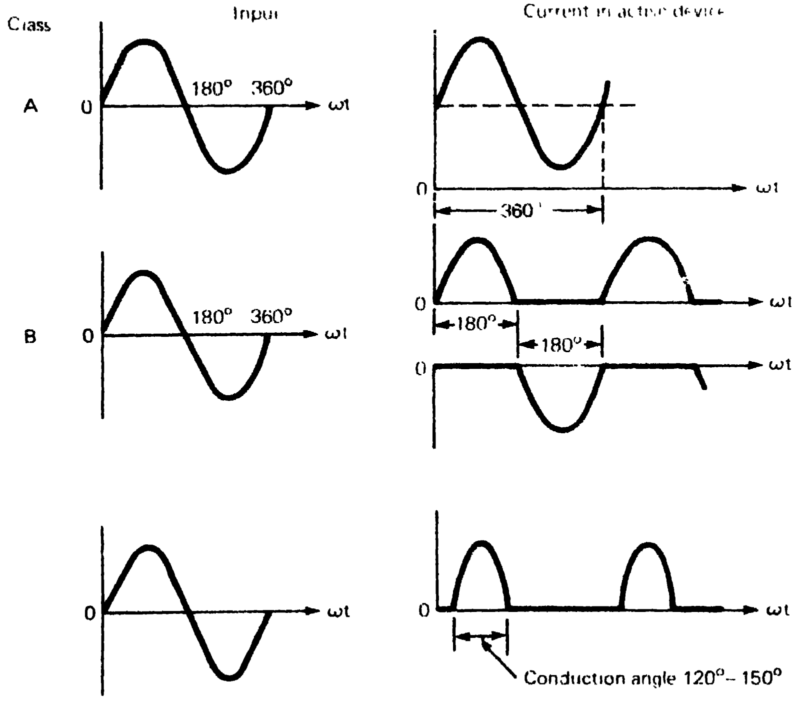


Fig. 7-7. Classes of amplifier operation.

In class A operation, current flows through the active device for the whole of the input cycle. In class B, two active devices are needed, with current flowing in each for only half of the input cycle. (An output transformer combines the two halves in a push-pull circuit to provide a full sine wave.) Class C operation has current flow in the active device for less than one half the input cycle (120° to 150°), so that only a non-sinusoidal pulse of current results. (This requires a tuned parallel resonant circuit as a load to provide a sinusoidal output, even if two push-pull active devices are used.)

Fig. 7-7 also indicates that only in class A does current flow through the active device when there is no input signal. There is little drain on the power supply in the absence of an input signal in classes B and C. There are also intermediate classes of operation, such as AB, but these will be covered in the particular applications.

7-11. HARMONIC DISTORTION

A good amplifier should not only produce an enlarged version at the output, but should also provide a faithful reproduction of the input waveform. One type of distortion that is common to most amplifiers is called amplitude or harmonic distortion, which results from the nonlinearity of the active device used for amplification. That is, the tube or semiconductor device may not increase equally all portions of the input signal over its positive and negative excursions.

The relation between the input signal (the excitation) and the output current in the load may be described by the dynamic transfer curve, an example of which is shown in Fig. 7-8. This is similar to the diode dynamic curve in Section 2-2, in which the output current waveform is obtained graphically from the input excitation waveform by projections intersecting on the transfer curve. It should be noted that the excitation may be a voltage waveform in tube and FET (field effect transistor) amplifiers but a current waveform in many bipolar transistor amplifiers.

In any event, if the excitation is sinusoidal, the output waveform will only be sinusoidal if the transfer curve is a straight line (indicating the active device is perfectly linear) or if only a small signal is applied. A portion of the curve may be sufficiently linear to cause no distortion in the output. Fig. 7-8 clearly shows how amplitude distortion results from a large signal input to an amplifier with a nonlinear dynamic transfer curve. Why this is called harmonic distortion is seen in the following discussion.

7-12. THREE-POINT METHOD OF CALCULATING HARMONIC DISTORTION

Let us assume that the input signal is large enough so that distortion results, but not so large that the amplifier is driven over the extremes of the dynamic transfer curve. In other words, the slight curvature of the dynamic curve over the region of operation may be described by a parabolic equation, such that the alternating portion of the output current, i_o, may be represented as a function of the excitation input, x, by the following:

$$i_o = K_1 x + K_2 x^2 \tag{7-14}$$

where K_1 and K_2 are constants. Let the excitation be sinusoidal and be described by

$$x = X_m \cos \omega t \tag{7-15}$$

where $\omega = 2\pi f$ is the frequency of the input signal.

Substitution of Eq. 7-15 in Eq. 7-14, with suitable trigonometric identities, yields the following equation for the total output current, i_O

$$i_O = I_Q + i_o \tag{7-16}$$

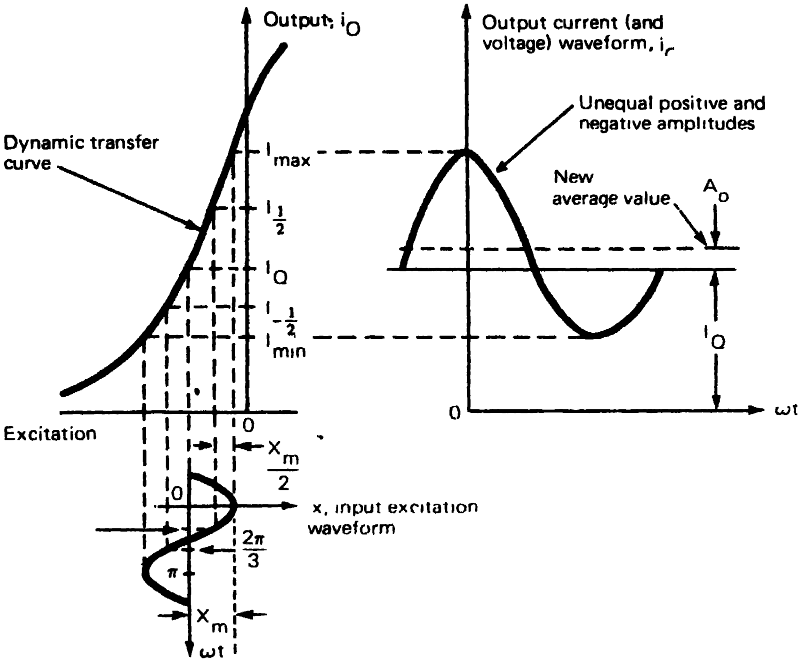

Output: i_O
Dynamic transfer curve
I_{max}
$I_{\frac{1}{2}}$
I_Q
$I_{-\frac{1}{2}}$
I_{min}
Excitation
0
X_m
$\frac{X_m}{2}$
x, input excitation waveform
$\frac{2\pi}{3}$
π
ωt
Output current (and voltage) waveform, i_r
Unequal positive and negative amplitudes
New average value
A_o
I_Q
0
ωt

Fig. 7-8. Graph of a dynamic transfer curve showing how distortion in the output results from a large input signal.

$$= I_Q + A_o + A_1 \cos \omega t + A_2 \cos 2\omega t \tag{7-17}$$

where

I_Q is the dc current with no signal (quiescent conditions).

A_o is an extra dc component due to rectification of the signal.

A_1 is the amplitude of the desired signal at the fundamental frequency, ω.

A_2 is the amplitude of the second harmonic frequency, 2ω.

The presence of an extra dc component A_o indicates a new average current through the active device. The slight shift in the operating point when a signal is applied usually indicates that distortion is taking place. This shift is shown in Fig. 7-8.

A_2 indicates that a component exists in the output that has twice the frequency of the desired input signal, which is termed the second harmonic. The degree of distortion is measured by the relative amount of second harmonic to the fundamental and is expressed as a percentage:

Percent 2nd harmonic distortion,

$$D_2 = \frac{|A_2|}{|A_1|} \times 100\% . \tag{7-18}$$

It remains to determine A_o, A_1 and A_2 in terms of the current quantities shown in Fig. 7-8. This may be done by considering the current i_O at three points, $\omega t = 0, \frac{\pi}{2}$, and π. Substituting these values of i_O in Eq. 7-17 results in three simultaneous equations, whose solution yields the following amplitudes:

$$A_1 = \frac{I_{max} - I_{min}}{2} \qquad (7\text{-}19)$$

$$A_o = A_2 = \frac{I_{max} + I_{min} - 2I_Q}{4} \qquad (7\text{-}20)$$

$$= \frac{I_{max} + I_{min}}{} - \frac{I_Q}{} . \qquad (7\text{-}21)$$

All of these current values may be obtained from the dynamic transfer curve, if drawn, or from the intersection of the load line drawn on the active device's characteristic curves, as will be shown later.

Note what A_o indicates. If the average value of I_{max} and I_{min},

$$\left(\text{equal to } \frac{(I_{max} + I_{min})}{2} \right),$$

equals I_Q then A_2 (and A_o) equals zero and there is no distortion. Any unequal variation in i_o above and below the no-signal average current will result in distortion. It is possible for A_2 to be negative, so D_2 is defined in terms of absolute values since a negative distortion is meaningless.

7-13. FIVE-POINT METHOD OF CALCULATING HARMONIC DISTORTION

In the three-point method, it was assumed that the input signal was not sufficiently large to drive the amplifier to the extremes of its dynamic transfer curve, and so the distortion was predominantly second harmonic in nature. If this is not true, harmonics higher than the second will not be negligible and may be determined as follows.

To represent the increased curvature toward the extremes of the dynamic curve, we now let i_o be represented by higher order terms than in Eq. 7-14:

$$i_o = K_1 x + K_2 x^2 + K_3 x^3 + K_4 x^4 + \ldots . \qquad (7\text{-}22)$$

Substitution of Eq. 7-15 in Eq. 7-22 yields the following equation for the total output current i_O:

$$i_O = I_Q + A_o + A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + A_4 \cos 4\omega t + \ldots . \qquad (7\text{-}23)$$

The notations are as in Section 7-12, but, in addition, note that harmonics higher than the second are now present. The amplitudes of harmonics up to the fourth may be obtained in a manner similar to the previous section, but now five values of current are necessary to solve the equations. These are chosen at equal intervals and include values at one-half the maximum positive amplitude of the input, $I_{\frac{1}{2}}$, and at one-half the maximum negative amplitude of the input, $I_{-\frac{1}{2}}$, as shown in Fig. 7-8. Using these designations, the amplitudes are

$$A_o = \frac{1}{6}(I_{max} + 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} + I_{min}) - I_Q \qquad (7\text{-}24)$$

$$A_1 = \frac{1}{3}(I_{max} + I_{\frac{1}{2}} - I_{-\frac{1}{2}} - I_{min}) \qquad (7\text{-}25)$$

$$A_2 = \frac{1}{4}(I_{max} - 2I_Q + I_{min}) \qquad (7\text{-}26)$$

$$A_3 = \frac{1}{6}(I_{max} - 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} - I_{min}) \qquad (7\text{-}27)$$

$$A_4 = \frac{1}{12}(I_{max} - 4I_{\frac{1}{2}} + 6I_Q - 4I_{-\frac{1}{2}} + I_{min}) . \qquad (7\text{-}28)$$

The individual contributions to harmonic distortion are given by

$$D_2 = \frac{|A_2|}{|A_1|} \times 100\% \qquad (7\text{-}29)$$

$$D_3 = \frac{|A_3|}{|A_1|} \times 100\% \qquad (7\text{-}30)$$

$$D_4 = \frac{|A_4|}{|A_1|} \times 100\% \tag{7-31}$$

and the total harmonic distortion by

$$D = \sqrt{D_2^{\,2} + D_3^{\,2} + D_4^{\,2} + \ldots} \tag{7-32}$$

Note that if distortion is present the calculation of output power at the fundamental frequency must be made using the amplitude, A_1:

$$P_1 = \frac{A_1^{\,2}}{2} R_L . \tag{7-33}$$

(Remember A_1 is the current amplitude with an rms value of $\frac{A_1}{\sqrt{2}}$.)

This will give a different value from using

$$P_1 = V_o I_o = \frac{(V_{max} - V_{min})}{2\sqrt{2}} \times \frac{(I_{max} - I_{min})}{2\sqrt{2}}$$

$$= \frac{(V_{max} - V_{min})(I_{max} - I_{min})}{8} , \tag{7-34}$$

since this assumes no distortion.

Theoretically the total output power is P_o

$$P_o = \frac{A_1^{\,2}}{2} R_L + \frac{A_2^{\,2}}{2} R_L + \frac{A_3^{\,2}}{2} R_L + \frac{A_4^{\,2}}{2} R_L \tag{7-35}$$

$$\frac{A_1^{\,2}}{2} \cdot R_L \left(1 + \frac{A_2^{\,2}}{A_1^{\,2}} + \frac{A_3^{\,2}}{A_1^{\,2}} + \frac{A_4^{\,2}}{A_1^{\,2}}\right) \tag{7-36}$$

$$= P_1 \left(1 + D_2^{\,2} + D_3^{\,2} + D_4^{\,2}\right) \tag{7-37}$$

$$= P_1 \,(1 + D^2) . \tag{7-38}$$

But if D is as much as 10%, P_o would be

$$P_o = P_1\,(1 + 0.1^2) = P_1\,(1.01) \approx P_1$$

so that the total power is essentially given by Eq. 7-33.

Numerical examples of the calculation of harmonic distortion for particular amplifiers are contained in Section 7-14 and the following chapters. Calculations are usually made with reference to the load line and the device's characteristic curves, rather than the dynamic transfer curve, but the notations used are the same. For high quality audio amplifiers the total harmonic distortion should not exceed 0.7 percent at the desired listening level, although distortion will of course increase at higher power levels as shown by the graph of harmonic distortion vs. frequency in Fig. 7-9. It should be stated here that the inherent nonlinearity of active devices is usually overcome in an amplifier system by the use of a technique called negative feedback, a topic to be fully explored in Chapter 17.

7-14. OSCILLOSCOPE DISPLAY OF AN AMPLIFIER'S DYNAMIC TRANSFER CURVE

A dynamic transfer curve relating an amplifier's output voltage to its input voltage may be displayed on an oscilloscope by using the method shown in Fig. 7-10. This may be recognized as the method of determining the phase angle by Lissajous figures. However, in this application, the frequency of the signal generator is adjusted to make the phase relationship between input and output either 0° or 180°, so that the trace on the oscilloscope screen is a single line, sloping either way, depending upon the phase and the oscilloscope. (A double line may still exist and may be removed (or reduced) by connecting the vertical input ahead of the output coupling capacitor, C_o.)

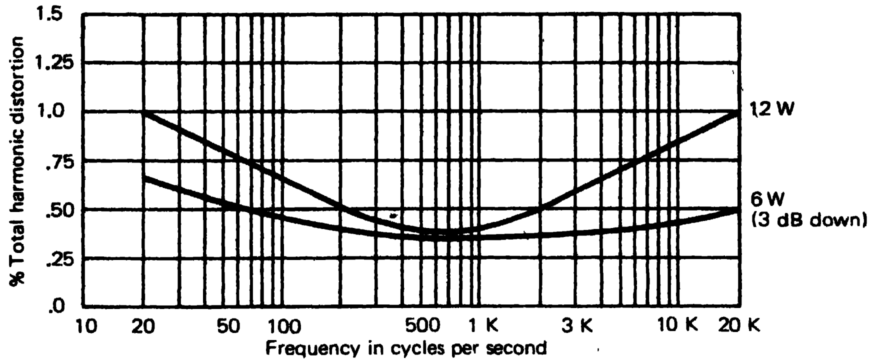


Fig. 7-9. Curve of harmonic distortion vs. frequency for a 12 watt audio amplifier.

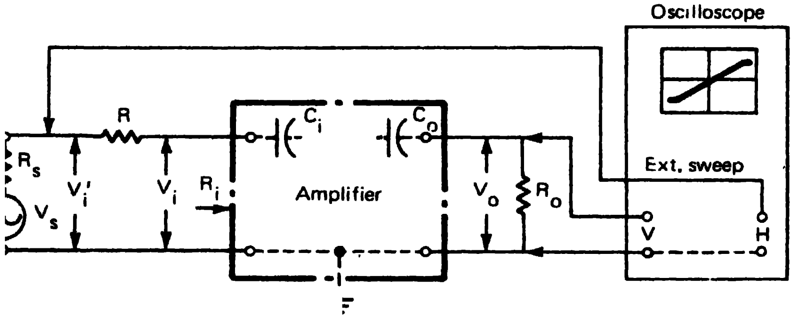


Fig. 7-10. Method of observing an amplifier's dynamic transfer curve.

Note that the input excitation voltage is applied to the horizontal amplifier. Since this often does not have enough gain to cause sufficient deflection on the oscilloscope if connected to V_i a large resistor R (compared with R_i) will provide a larger signal at V_i'. The input signal is increased until curvature at the extremes of the dynamic curve is observed, indicating distortion of the output.

Using the vertical and horizontal calibrations of the oscilloscope yields the dynamic transfer curve with respect to the origin (the center of the screen). The dc-shifts (upward and to the left,) to give a curve that conforms to Fig. 7-11, may be obtained by measuring the dc levels across the input and output coupling capacitors C_i and C_o.

The output voltage is proportional to the output current. Thus, a calculation of the total harmonic distortion may be made from this curve, (for a given input voltage), using the equations in Section 7-12 or 7-13, with V_o replacing I, as shown in the following example.

Example 7-6. An amplifier's dynamic transfer curve, obtained from an oscilloscope display, is shown in Fig. 7-11.

(a) Determine the percent second harmonic distortion for an input signal of 0.4 V peak-to-peak.

(b) Determine the total harmonic distortion for an input signal of 0.8 V peak-to-peak.

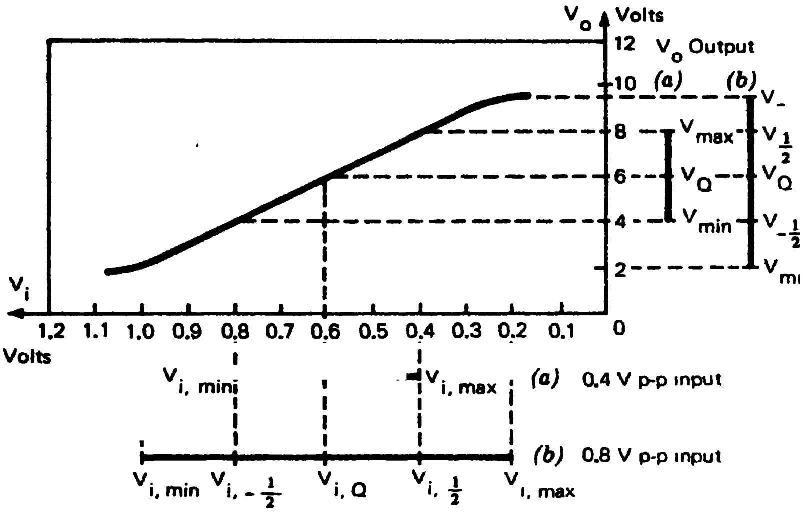


Fig. 7-11. Dynamic transfer curve to determine the harmonic distortion for 0.4 V and 0.8 V peak-to-peak inputs.

Solution

(a) 0.4 V peak-to-peak input:

$V_Q = 6\text{ V},\ V_{max} = 8\text{ V},\ V_{min} = 4\text{ V}.$

Since the input signal does not drive the amplifier into the extreme curvature regions the three-point method may be used:

$$A_1 = \frac{V_{max} - V_{min}}{2} = \frac{8-4}{2} = 2\text{ V}$$

$$A_o = A_2 = \frac{V_{max} + V_{min} - 2V_Q}{4} = \frac{8+4-12}{4} = 0.$$

This indicates that there is no distortion in the output and that there will be no change in dc current or voltage levels. A measurement of the total harmonic distortion with an H-P 332A distortion analyzer indicated 0.9% in V_o. However, the input to the amplifier had a distortion of 0.6%, so the distortion contributed by the amplifier is approximately only 0.7%, a value too low to show up with the scale used in Fig. 7-11. Also, there was no detected change in current or voltage level.

(b) 0.8 V peak-to-peak input:

This requires that the five-point method be used since we are now operating on the curved portion of the dynamic transfer curve, and the higher harmonics are not negligible.

$V_{max} = 9.5\text{ V},\ V_{\frac{1}{2}} = 8\text{ V},\ V_Q = 6\text{ V},\ V_{-\frac{1}{2}} = 4\text{ V},$

$V_{min} = 2\text{ V}.$

Substituting these values in the equations of Section 7-13 (with V replacing I) yields the following:

$A_o = -0.09\text{ V}$ $\qquad D_2 = 3.26\%$

$A_1 = 3.83\text{ V}$ $\qquad D_3 = 2.17\%$

$A_2 = -0.125\text{ V}$ $\qquad D_4 = 1.07\%$

$A_3 = -0.083\text{ V}$

$A_4 = -0.041\text{ V}$

Total harmonic distortion $D = 4.1\%$.

This compares with a measured value of 3.8%. An increase in dc current into the amplifier of 0.15 mA was observed when the input was applied. This current flowing through a 560-ohm load resistor (in the amplifier) caused a decrease of 0.084 V in the output level, which compares well with the calculated value of -0.09 V.

7-15. MEASUREMENT OF HARMONIC DISTORTION

Total harmonic distortion may be measured directly in percent using a distortion analyzer, a simplified block diagram of which is shown in Fig. 7-12.

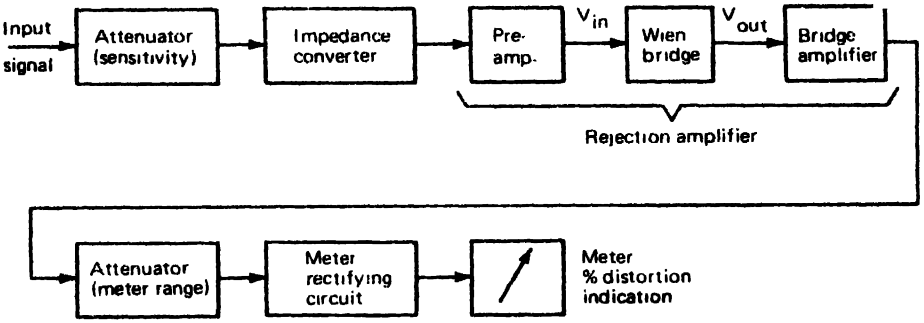

Input signal
Attenuator (sensitivity)
Impedance converter
Pre-amp.
V_{in}
Wien bridge
V_{out}
Bridge amplifier
Rejection amplifier
Attenuator (meter range)
Meter rectifying circuit
Meter % distortion indication

Fig. 7-12. Simplified block diagram of a distortion analyzer.

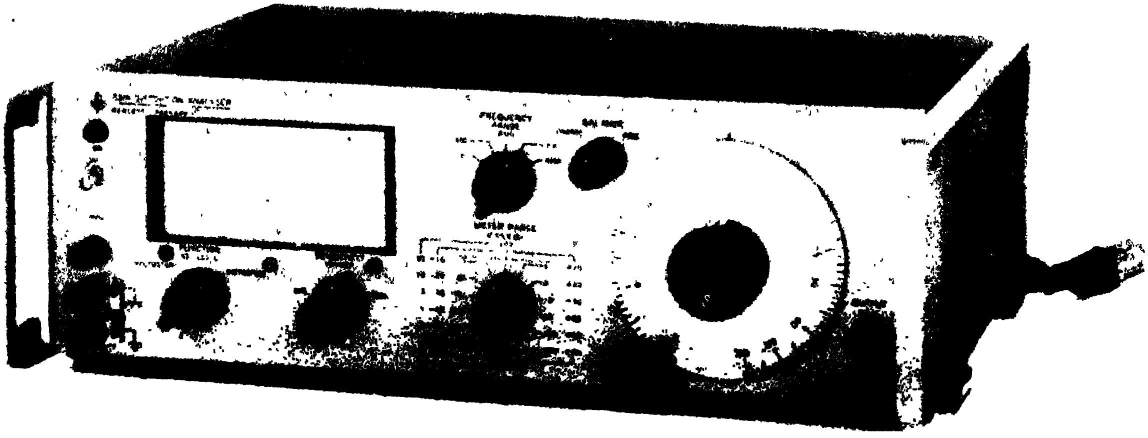


Fig. 7-13. A commercial distortion analyzer. (Courtesy of Hewlett-Packard Co.)

The input attenuator (sensitivity control) along with the meter range attenuator accommodates input signals from 0.3 V to 300 V typical. The impedance converter provides a high input impedance so that its output voltage is independent of the internal resistance of the source whose voltage is being analyzed for distortion. The heart of the distortion analyzer is the rejection amplifier which consists basically of a Wien bridge that may be tuned to the fundamental frequency of the voltage being analyzed. This is done by adjustment of the components shown in Fig. 7-14(a).

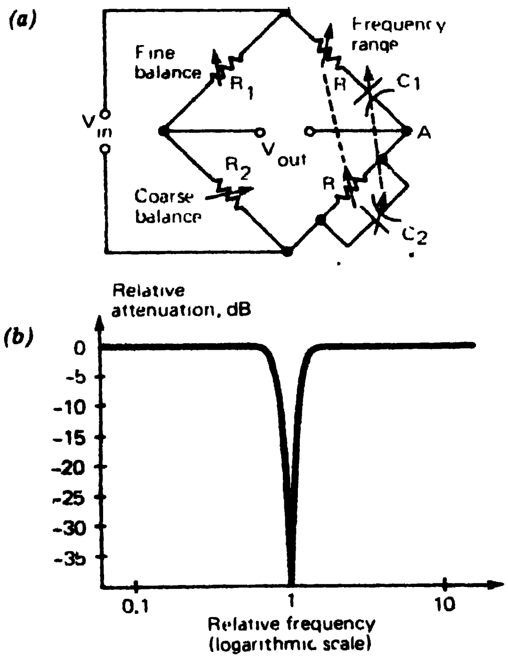


Fig. 7-14. (a) Typical Wien bridge showing components to be adjusted for balance.
(b) Rejection characteristics of the Wien bridge.

The Wien bridge is a sinusoidal frequency-selective circuit, which, when tuned to a given frequency by RC_1 and RC_2 and correctly balanced by R_1, R_2, will amost completely block the passage of a pure sine wave from V_{in} to V_{out}.

If V_{in} has a fundamental frequency the same as that to which the bridge has been tuned and some harmonics of distortion, the fundamental will be rejected, and only the distortion components will appear at V_{out}. That is, the bridge will not attenuate frequencies at two or three times that of the fundamental, as shown in the rejection characteristics of Fig. 7-14(b). This very narrow cone of rejection is achieved by feedback in the amplifiers before and after the Wien bridge.

The voltage appearing at V_{out} is rectified and is used to provide a visual readout on a meter calibrated in percent distortion. Adjustment is made for 100% distortion indication when the instrument is placed in the "set-level" mode. In this position, point A of the Wien bridge is connected to ground, which disables the frequency rejecting capability of the bridge and allows the total value of V_{in} (fundamental and harmonics) to cause meter deflection. Vernier adjustment on the sensitivity control is made to give full scale deflection representing 100% distortion.

A Hewlett Packard 331A solid-state distortion analyzer pictured in Fig. 7-13 can measure distortion levels from 0.1% to 100% full scale deflection on seven ranges for frequencies from 5 Hz to 600 kHz. Provision is also made for observing the distortion on an oscilloscope.

7-16. OTHER TYPES OF AMPLIFIER DISTORTION

Intermodulation Distortion

Another source of distortion in amplifiers arises when the input signal consists of two or more frequencies f_1 and f_2 (which is usually the case, since most speech or music consists of complex waveforms that have many different frequencies). If the active device is nonlinear, the output will consist of the fundamental; the harmonic components, and, in addition, components at frequencies of $f_1 + f_2$ and $f_1 - f_2$. These components at the sum and difference frequencies are called intermodulation frequencies and detract from the original intelligence. Intermodulation (IM) distortion may be expressed as a percentage of the fundamental and should not exceed 2% in a high quality audio amplifier. A typical curve of IM distortion vs. power output is shown in Fig. 7-15 for a 12-watt audio amplifier.

Frequency Distortion

Over the range of frequencies at which it is to be used, an amplifier should ideally provide the same amplification for all frequencies. However, generally at the lower and higher extremes of the range there is a loss in gain, which is usually due to the capacitive components in the amplifier as well as the active device itself. This is usually illustrated in a frequency response curve, which will be covered in detail in the next chapter, after a discussion of decibels.

Phase Distortion

There is usually some phase shift between the input and output waveforms, and in audio amplifiers the shift may vary with frequency. This is called phase distortion, but, since the ear is not sensitive to phase shift, it is not objectionable in an audio amplifier. However, in video amplifiers, a change in phase shift with frequency can be tolerated only if the angle of phase shift is proportional to frequency, so that the time delay is constant. If the phase shift varies with frequency in any other way, different frequency components of the signal are delayed by different amounts of time, and so, for example, a television picture is "smeared" by what is often referred to as *delay distortion*. The measurement and graphical presentation of phase shift will be covered in the next chapter.

7-17. SUMMARY

(1) An ac amplifier contains one or more active devices and converts energy from the dc supply into ac power at the output that is proportional to the input signal in amplitude and frequency.

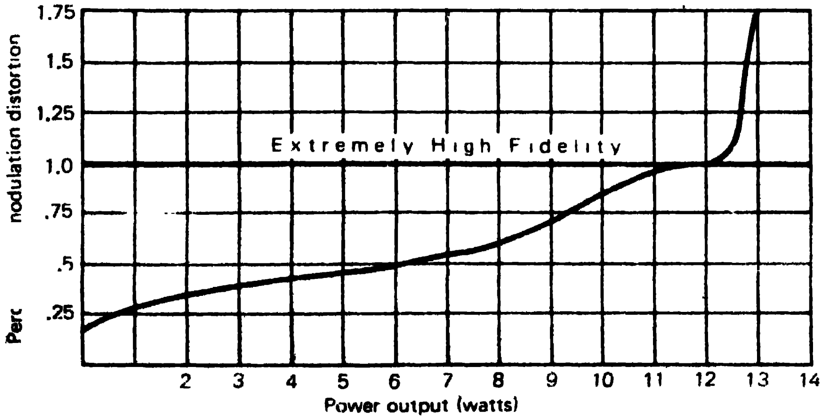


Fig. 7-15. Graph of intermodulation distortion vs. output power for a 12-watt amplifier.

(2) Current gain, voltage gain, and power gain, given by

$$A_i = \frac{I_o}{I_s},\ A_v = \frac{V_o}{V_i},\ A_p = A_i A_v,$$

are measures of the amplification in an amplifier between input and output.

(3) The amplifier's input resistance R_i may be determined from $R_i = \frac{V_i}{I_s}$, and its ratio to the source resistance determines the necessary value of open circuit source voltage for a required input to the amplifier.

(4) An amplifier's output resistance R_o' causes the voltage gain of the amplifier to be different with different values of load resistance R_o.

(5) An ideal amplifier has an infinite input resistance and zero output resistance.

(6) The conversion efficiency of a single-stage amplifier is the ac signal power delivered to the total load, divided by the dc input power to the active device.

(7) The three main classes of amplifier operation are A, B and C, in which current through the active device flows for the whole of the input cycle, half of the input cycle or less than half of the input cycle, respectively.

(8) Harmonic distortion is caused by the nonlinearity of an active device's characteristics and may be observed in an amplifier's dynamic transfer curve, which is a plot of output voltage or current against input excitation.

(9) Second harmonic distortion may be calculated from the dynamic transfer curve using the three-point method for small signals, or the total harmonic distortion may be obtained for large signals by the five-point method.

(10) Other types of amplifier distortion include: intermodulation distortion, which is the generation of components at the sum and difference frequencies; frequency distortion, in which signals at different frequencies are amplified by different amounts; and phase distortion, where the phase shift between input and output depends upon the signal frequency.

QUESTIONS

Q7-1. Explain how an ac amplifier may be thought of as an energy conversion circuit.

Q7-2. When calculating the voltage gain of an amplifier, is it necessary to convert all voltages to rms values? Explain.

Q7-3. (a) Why is it generally beneficial for an amplifier to have a high input resistance?
(b) Under what conditions is the input voltage to an amplifier substantially the same as the open circuit source voltage?
(c) Would the input resistance to an amplifier be a factor to take into consideration if all signal sources had zero internal resistance?

Q7-4. (a) What would be the output resistance of an ideal amplifier? Explain.
(b) Describe an experimental method for determining the output resistance of an amplifier.

Q7-5. How does the load resistance connected to an amplifier for maximum power transfer differ from the resistance used if voltage gain is of primary importance?

Q7-6. Explain how the conversion efficiency of an amplifier determines the power drain on the supply.

Q7-7. Which class of amplifier operation causes a lower drain on the power supply when no signal is being handled? Explain.

Q7-8. Under what conditions will a sinusoidal input to an amplifier provide a sinusoidal output with no distortion?

Q7-9. What does a change in the average current through an active device indicate when a signal is applied to a class A amplifier?

Q7-10. Why is nonlinear distortion referred to as harmonic distortion?

Q7-11. When is it necessary to use the five-point method to calculate harmonic distortion?

Q7-12. (a) What is meant by the notation $I_{\frac{1}{2}}$ and $I_{-\frac{1}{2}}$?
(b) Is the input excitation necessarily a voltage?

Q7-13. Explain why an amplifier's harmonic distortion will increase at higher power levels.

Q7-14. Explain how the Wien bridge with feedback is the heart of the harmonic distortion analyzer.

Q7-15. What are considered to be the maximum acceptable levels of

(a) total harmonic distortion
(b) intermodulation distortion, in a high quality audio amplifier?

Q7-16. (a) What other types of distortion exist in an amplifier besides those mentioned in Q7-15?
(b) Are they both objectionable in an audio amplifier?

PROBLEMS

P7-1. An amplifier has an input signal of 5 V at 0.1 mA. The output of the amplifier provides 10 mA to a 450-ohm load. Determine:

(a) The current, voltage and power gains.
(b) The input resistance to the amplifier.

P7-2. The input voltage to an audio amplifier is 25 mV. An oscilloscope connected across a 1 M ohm resistor in series with the input to the amplifier indicates 1 V peak-to-peak. If the voltage across a 10-K ohm load is 3.5 V, determine:

(a) The current, voltage and power gains.
(b) The amplifier's input resistance.

P7-3. (a) What will be the input voltage to the amplifier in P7-1 when a 9 V source with a 25-K ohm internal resistance is connected to the amplifier?
(b) What will be the amplifier's output voltage?

P7-4. (a) What must be the open circuit source voltage to provide 0.5 V peak-to-peak at an amplifier's input terminals if the amplifier has an input resistance of 5 K ohms, for the following values of source resistance:
(i) 600 ohms, (ii) 0 ohms, (iii) 5 K ohms?

(b) Assuming the amplifier has a constant voltage gain of 10, what will be the output voltage at each value of source resistance?

P7-5. A given amplifier is capable of presenting three different input resistances of 200 ohms, 50 K ohms, and 1 M ohm. If the amplifier's power gain is 1000 and its voltage gain is 100, determine which input to use to provide maximum power output and which to use to provide maximum output voltage for a low impedance microphone of 500 ohms and 20 mV open circuit voltage. Determine the power and voltage outputs.

P7-6. An amplifier's open circuit output voltage is 10 V peak-to-peak. If, when a 10-K ohm load is connected, the output voltage drops to 8 V peak-to-peak, determine the amplifier's output resistance.

P7-7. What will be the output voltage of the amplifier in P7-6 when the following loads are connected:

(a) 2.5 K? (b) 1 K ohm? (c) 1 M ohm?

P7-8. A 9 V transistor radio playing at full volume draws 200 mA to provide 1/2 watt of signal power to the loudspeaker. Determine the approximate conversion efficiency.

P7-9. If a 9 V battery has a 5 ampere-hour rating, determine approximately how long a radio will play when delivering 100 mW at a conversion efficiency of 40%.

P7-10. Substitute Eq. 7-15 into 7-14, and with the use of

$$\cos^2 \omega t = \frac{1 + \cos 2\omega t}{2}$$

determine the coefficients A_o, A_1 and A_2 in Eq. 7-17, in terms of K_1, K_2, and X_m.

P7-11. A class A audio amplifier, under quiescent conditions (no signal applied), has a dc current demand of 25 mA. This current is observed to increase to 26 mA when a signal is applied to provide a 40 V p-p output across a 1 K ohm load connected to the amplifier. Determine:

(a) the amplitude of the fundamental.
(b) the amplitude of the second harmonic and the second harmonic distortion.
(c) the maximum and minimum values of output current.
(d) the output power at the fundamental frequency.
(e) the total output power.

Assume the 3-point method of determining distortion is sufficiently accurate.

P7-12. Using the dynamic transfer curve of Fig. 7-11, calculate:

(a) the percent second harmonic distortion in the output for a 0.6 V peak-to-peak input, with the no-signal operating point $V_{i,Q}$ at 0.6 V.
(b) the percent second harmonic distortion in the output for a 0.6 V peak-to-peak input, with the no-signal operating point $V_{i,Q}$ at 0.5 V.

P7-13. Using the dynamic transfer curve of Fig. 7-11, calculate:

(a) The total harmonic distortion in the output for a 0.9 V peak-to-peak input, $V_{i,Q}$ at 0.6 V.
(b) Assuming the output from part (a) is across a 1 K ohm load, determine the output power at the fundamental frequency using Eq. 7-33 and then by use of Eq. 7-34. Also find the total output power using 7-38 and compare with that from 7-33.

P7-14. The total output power from an amplifier is indicated on an output power meter to be 10.2 watts, but a calculation of the output power at the fundamental frequency gives 10 watts. What is the total harmonic distortion based on these two figures?

8

GENERAL AMPLIFIER CHARACTERISTICS II: DECIBELS

8-0. INTRODUCTION

The topic of decibels is very important in the field of electronics and sound measurement. It is an increasingly used word in this day of noise pollution, and one must be very careful in stating a noise level in decibels, not to make some of the common errors. Decibels are used to describe an amplifier's frequency response as well as the variation that occurs when bass and treble tone controls are used in an amplifier. These topics along with the square-wave testing of an audio amplifier are covered in this chapter.

8-1. DECIBELS

The reaction of the individual, either in the field of seeing or hearing, is nonlinear. At night for example, the glare of a headlight can be almost blinding, yet in midday is hardly noticeable in the presence of the sun. Two headlights would double the effect on vision at night but it would take two suns to double the effect of brightness during the day. That is, a doubling effect takes place in both cases, but a much greater increase is necessary in the latter case. The same is true in music, where a doubling of frequency from 128 Hz to 256 Hz (an increase of one octave) gives approximately the same *apparent change* in pitch as a change in frequency from 1024 to 2048 Hz (an increase of one octave).

To linearize this scale, a mathematical ratio is used. This is done by taking the logarithm to the base 10 of the ratio of the two power levels involved:

$$\text{Number of bels} = \log_{10} \frac{P_2}{P_1}, \tag{8-1}$$

where P_1 is the reference level that P_2 is being compared with. Since the bel (used to honor Alexander Graham Bell) is a rather large ratio, a smaller value called a decibel (one tenth of a bel) is commonly used:

$$\text{Number of decibels} = 10 \log_{10} \frac{P_2}{P_1}. \tag{8-2}$$

The word decibel is usually abbreviated to dB.

For example, if an amplifier has its power output increased from 4 watts to 8 watts, the increase measured in decibels, dB, is

$$\begin{aligned} 10 \log_{10} \left(\frac{8 \text{ watts}}{4 \text{ watts}}\right) &= 10 \log 2 \\ &= 10 \times 0.3010 = 3.01 \text{ dB}. \end{aligned}$$

This is an increase of +3 dB for an increase of 4 watts from an initial level of 4 watts.

Now consider the power being increased from 8 watts to 16 watts. This is an increase of

$$\begin{aligned} 10 \log \left(\frac{16 \text{ watts}}{8 \text{ watts}}\right) &= 10 \log 2 \\ &= 10 \times 0.3010 = 3 \text{ dB}. \end{aligned}$$

This is also an increase of 3 dB, but for an increase of 8 watts from an initial level of 8 watts.

Since the doubling of power will sound to be the same increase in each case (even though one is an increase of 8 watts and the other is an increase of 4 watts), the mathematical description of 3 dB in each case has indicated the same change in hearing response regardless of level. (This assumes the same frequency is being used in each case, since the ear has a sensitivity that varies with frequency as discussed in Section 8-12.) This is now a linear response.

A change of 3 dB can be noticed but is not an appreciable change in power level.

Example 8-1. Determine the power gain in decibels of a power amplifier that has an input voltage V_i of 15 V peak-to-peak and an output voltage V_o of 8 V peak-to-peak across a load resistance R_o of 4 ohms. The amplifier's input resistance is 470 K ohms.

Solution. The input power from the source is

$$= \frac{V_i^2}{R_i} = \frac{\left(\frac{V_{i,p\text{-}p}}{2\sqrt{2}}\right)^2}{R_i} = \frac{\frac{15^2}{8}}{470 \times 10^3}\ \text{W} = 60\ \mu\text{W}$$

$$P_o = \frac{V_o^2}{R_o} = \frac{\left(\frac{V_{o,p\text{-}p}}{2\sqrt{2}}\right)^2}{R_o} = \frac{\left(\frac{8}{2\sqrt{2}}\right)^2}{4} = 2\ \text{W}.$$

Therefore, the power amplifier has a power gain ratio of

$$A_p = \frac{P_o}{P_i} = \frac{2\ \text{W}}{60 \times 10^{-6}\ \text{W}} = \frac{2000 \times 10^3}{60} = 33.3 \times 10^3$$

To present this on a linear scale, this would be a decibel gain of

$$\begin{aligned} 10 \log\left(\frac{P_o}{P_i}\right) &= 10 \log (33.3 \times 10^3) \\ &= 10 \log (3.33 \times 10^4) \\ &= 10\,(4.523) = 45.23\ \text{dB}. \end{aligned}$$

8-2. OTHER EQUATIONS FOR DECIBEL COMPUTATION

Although the decibel was defined as a *power* ratio, it can also be used with voltage levels provided certain precautions are taken, such as knowing what impedance the voltages were measured across. That is, since

$$P = \frac{V^2}{R}$$

the equation for dB gain can be rewritten:

$$\text{dB} = 10 \log\left(\frac{P_o}{P_i}\right) = 10 \log\left(\frac{\frac{V_o^2}{R_o}}{\frac{V_i^2}{R_i}}\right)$$

$$= 10 \log\left(\frac{V_o^2}{V_i^2} \cdot \frac{R_i}{R_o}\right)$$

$$= 10\left[\log\left(\frac{V_o}{V_i}\right)^2 + \log\left(\frac{R_i}{R_o}\right)\right]$$

$$= 10 \log\left(\frac{V_o}{V_i}\right)^2 + 10 \log \frac{R_i}{R_o}$$

$$\therefore \text{dB} = 20 \log \frac{V_o}{V_i} + 10 \log \frac{R_i}{R_o}. \qquad (8\text{-}3)$$

Alternatively, it may be shown that

$$\text{dB} = 20 \log\left(\frac{I_o}{I_i}\right) + 10 \log\left(\frac{R_o}{R_i}\right), \qquad (8\text{-}4)$$

where I_i and I_o are input and output currents respectively. Note that $\text{dB} = 20 \log\left(\frac{V_o}{V_i}\right)$ only when these voltages are measured across the same value of resistance, $R_i = R_o$. That is, $10 \log 1 = 0$.

Example 8-2. Determine the power gain in Example 8-1 using Eq. 8-3.

Solution

$$\text{dB} = 20 \log\left(\frac{V_o}{V_i}\right) + 10 \log\left(\frac{R_i}{R_o}\right)$$

$$= 20 \log\left(\frac{8}{15}\right) + 10 \log\left(\frac{470 \times 10^3}{4}\right)$$

$$= -20 \log\left(\frac{15}{8}\right) + 10 \log (1.175 \times 10^5)$$

$$= -20 \log 1.875 + 10 \log (1.175 \times 10^5)$$

$$= -20 \times 0.273 + 10 \times 5.070$$

$$= -5.46 + 50.70 = 45.24 \text{ dB},$$ the same as before.

Note that, had the impedance correction factor not been included, the amplifier gain would have been −5.46 dB or a loss of +5.46 dB. To properly understand the use of the negative and positive signs for decibels, consider the following example.

Example 8-3. If the input to a network was 12 V and the output 6 V, and the input and output impedances were equal, determine the dB *gain* across the network.

Solution

$$\text{Gain in dB} = 20 \log \frac{V_o}{V_i} = 20 \log \frac{6}{12}$$

$$= 20 \log 0.5 = 20 \log 5 \times 10^{-1}$$

$$= 20 (\log 5 + \log 10^{-1})$$

$$= 20 (0.7 + (-1))$$

$$= 20 (-0.3) = -6$$

Note that taking the log of a fraction had to be done carefully. Since the log of a number less than 1 is always negative, an alternative method is to invert the fraction immediately and make the result negative as follows:

$$\text{dB} = 20 \log \frac{6}{12} = -20 \log \frac{12}{6}$$

$$= -20 \log 2 = -20 \times 0.3 = -6 \text{ dB}.$$

In either case, the gain is −6 dB, or a loss of +6 dB. Note, we do *not* say a loss of −6 dB. The minus sign in the second case gives the effect of a double negative and would not be correct.

8-3. ZERO DECIBEL REFERENCE LEVEL

The preceding discussion has considered the *increase* in power from one level to another. We did not refer to any absolute level of power in decibels; the increase was always relative to the initial level. In order to have an absolute level of power in decibels, one must specify what is meant by a zero decibel reference. Note that 0 watts cannot be used as zero dB. This is shown by the following:

$$\text{dB} = 10 \log \frac{P_o}{P_i} = 10 \log \frac{P_o}{0} = 10 \log \quad = \infty ?$$

That is, any power level compared with zero power is infinitely larger. It is difficult to take the logarithm of infinity! Thus decibels would not be defined if zero power were used as the zero reference level.

There are three zero references in common use, depending upon the industry involved:

Electronics Industry Association (EIA): The zero decibel reference is a 6-mW power dissipated in a 500-ohm resistive load.

Radio Broadcasting Industry Association (RIA): The zero decibel reference is a 1-mW power dissipated in a 600-ohm resistive load. This is often referred to as dBm, to indicate it uses one milliwatt as a reference and is different from the EIA reference.

The Volume Unit (VU): The volume unit reference is a 1-mW power dissipated in a 600-ohm resistive load. If a meter calibrated in VU is used on pure sine waves at 1000 Hz, the readings will be numerically equal to dBm. However, a VU meter is generally used to indicate complex waves of speech and music, and the units are not in decibels. To obtain decibels, a correction factor must be subtracted from the VU readings. A VU meter is the type used on communications receivers, tuners, and tape recorders, where the power level varies quite rapidly and requires a high degree of damping in the meter. The meter is usually calibrated from −20 to +3 VU and gives an indication of the volume of the complex signal.

A decibel meter is used in applications where the power level is not changing rapidly, and is lightly damped. It is generally used with signals at a specific test frequency.

8-4. USE OF A VOLTMETER AS A DECIBEL INDICATOR

It is to be noted that the zero decibel reference levels referred to in the previous section all specify a resistance value as well as the power level. This is done so that voltmeters may be calibrated in decibels, but their scales in dB are only correct if the voltmeter is being used across a resistance equal to the value for which the voltmeter was calibrated. That is, if a voltmeter is being used to measure 6 V across 6 ohms this is definitely a different power level from when 6 V are measured across 600 ohms. In the latter case the power is

$$\frac{V^2}{R} = \frac{6^2}{600} = 60\ \mathrm{mW},$$

whereas in the first case it is $\frac{6^2}{6} = 6$ W. Therefore, it is not simply a matter of putting a dB scale on the voltmeter scale and reading off dB directly. It would depend upon what reference level were used for 0 dB.

In much electronic work and especially in communications, the standard reference in use is the RIA level, 1 mW in 600 ohms. Consider what voltage this is:

$$P = \frac{V^2}{R} \qquad V^2 = RP = 600 \times 1 \times 10^{-3} = 0.6$$

$$\therefore V = \sqrt{0.6} = 0.775\ \mathrm{V}.$$

Thus, if a voltmeter on the 1.5-volt scale were placed across a 600-ohm resistor that had 1 mW dissipated in it, the deflection on the 1.5-volt scale to 0.775 volt would correspond to 0 dB(m). Say the voltage increased to 1.1 volts. What dB level should this represent?

$$\text{New power} = \frac{V^2}{R} = \frac{1.1^2}{600} = \frac{1.21}{600} = 2\ \mathrm{mW}.$$

$$\mathrm{dB} = 10 \log\left(\frac{P_o}{P_i}\right) = 10 \log\left(\frac{2\ \mathrm{mW}}{1\ \mathrm{mW}}\right)$$

$$= 10 \log 2 = 10 \times 0.3 = 3\ \mathrm{dB}.$$

1.1 V on the 1.5 V scale is equivalent to +3 dB. Similarly, if instead of the voltage increasing by a factor of $\sqrt{2}$ from 0.775 V to 1.1 V, it decreased from 0.775 to $\frac{0.775}{\sqrt{2}} = 0.548$ V, the new power level would be only 1/2 mW, and the dB level would be -3 dB, so 0.548 V on the 1.5 V scale should be marked -3 dB. Thus, the 1.5 V scale can be calibrated from -10 dB to +5 dB, as indicated in Fig. 8-1.

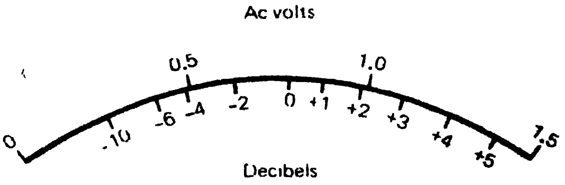


Fig. 8-1. Voltmeter scale calibrated in decibels for 0 dB = 1 mW in 600 ohms.

8-5. VOLTMETER RANGE CORRECTION FACTORS

If in a particular situation it is required to know what dB level exists across a 600 ohm resistor when the voltage is, say, 12 V, switch to the 15-V range. The needle will line up with +3.8 dB on the 1.5-V scale. However, a correction factor must be added since the range is 10 times larger. This correction is:

$$20 \log\left(\frac{V_o}{V_i}\right) = 20 \log\left(\frac{15\ \mathrm{V}}{1.5\ \mathrm{V}}\right) = 20 \log 10$$

$$= 20 \times 1 = 20\ \mathrm{dB}.$$

Therefore the true level is 3.8 + 20 = 23.8 dB.

If, in a particular application, on the 15 V range, the needle lined up with -2 dB (approximately 6.2 V), then adding the correction factor of 20 dB would result in -2 + 20 = 18 dB.

The correction factors to be added are shown in Table 8-1. These are true for the Heathkit IMW-11 VTVM, that uses 1 mW in 600 ohms as 0 dB, and assumes that readings are taken across a 600-ohm load.

TABLE 8-1. VOLTMETER RANGE CORRECTION FACTORS

AC VOLTS RANGE	DECIBEL SCALE
0-1.5 volts	Read dB directly
0-5 volts	Add 10 dB to the reading
0-15 volts	Add 20 dB to the reading
0-50 volts	Add 30 dB to the reading
0 150 volts	Add 40 dB to the reading
0-500 volts	Add 50 dB to the reading
0 1500 volts	Add 60 dB to the reading

8-6. IMPEDANCE CORRECTION FACTOR

If the voltage is being read across a resistor of say 60 ohms instead of 600 ohms, an additional *impedance* correction factor is necessary to give *true* decibels above the 0 dB reference level. Since a given voltage across 60 ohms is a larger power than across 600 ohms this is a positive amount and is

$$10 \log \left(\frac{R_2}{R_1}\right) = 10 \log \frac{600}{60} = 10 \log 10$$

$$= 10 \times 1 = 10 \text{ dB}.$$

To take the last example above (approximately 6.2 V across 60 ohms), the true dB reading would be -2 + 20 + 10 = 28 dB (direct reading = -2 dB, range correction = +20 dB, impedance correction = +10 dB). Note that, had the 6.2 V been measured across 6000 ohms, the impedance correction factor would be negative = -10 dB, and the true power level would be -2 + 20 - 10 = 8 dB.

These corrections are necessary since the voltmeter reacts to voltage levels and not power levels. The voltmeter is not aware of what resistance the voltage is across, and so this correction must be made by the operator. For a voltmeter calibrated in decibels based on a reference level of 0 dB = 1 mW in 600 ohms we can generalize the expression for true decibels as follows:

True dB = Direct dB reading + Range correction + $10 \log \frac{600}{R_o}$, (8-5)

where R_o is the load resistance in ohms. Note that if R_o is less than 600 ohms, the impedance correction factor is $+ 10 \log \left(\frac{600}{R_o}\right)$, or, if R_o is more than 600 ohms, the correction factor is more easily obtained by using $-10 \log \frac{R_o}{600}$, to avoid taking the logarithm of a decimal number.

If a voltmeter is calibrated for use across a 500-ohm resistance, the impedance correction factor would have 500 substituted for 600 in the above expressions.

8-7. FREQUENCY RESPONSE CURVES

Now that we have covered the topic of decibels we can apply this information to another very important general amplifier characteristic, its frequency response. Over the range of frequencies at which it is to be used, an amplifier should ideally provide the same amplification for all frequencies. The degree to which this is done is usually indicated by a frequency response curve, which is a plot of the amplifier's output voltage or power against frequency of input signal. However, to show the relative "flatness" of the amplifier's gain with frequency, the response curve is often plotted directly in dB with the gain at 1 kHz (or 400 Hz) taken as the reference level of zero dB and all other gains indicated as positive or negative dB with respect to this reference. This can be done using a voltmeter having a dB scale, regardless of what resistance the voltage is measured across, in the following manner.

With the amplifier delivering its rated output into a dummy load resistance, connect the voltmeter across the resistance and switch to a range that can accommodate the voltage at the output, such that the pointer deflection is near midscale. Then adjust the input to the amplifier at 1 kHz until the pointer on the voltmeter lines up with 0 dB on the dB scale. Keeping the amplitude of the signal to the amplifier constant, vary the frequency of the generator, noting the dB level at each desired frequency by reading directly off the dB scale, either positive or negative. These are the values that would be plotted on the graph paper as shown in Fig. 8-2.

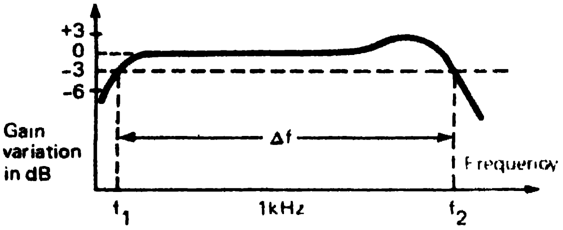


Fig. 8-2. Typical frequency response curve of an audio amplifier.

This curve shows a resonant rise that may result at higher frequencies due to the distributed capacitance and inductance of the output transformer.

In order to present the frequency response curve on a convenient length of paper the large variation in frequency from 10 Hz to 100 kHz is drawn on a logarithmic scale. The vertical scale in decibels uses a linear scale so the paper used for frequency response curves is referred to as semilog paper, a 5-cycle example of which is shown in Fig. 8-3 for a commercial 12-watt amplifier. If it is required that the above curves give the actual power level in decibels (not the relative power level), a correction factor must be added, which must take into account both the voltage range that was used and the impedance correction.

It is interesting to note that, for the public, audio amplifier frequency response curves are drawn with the frequency indicated in hertz or cycles per second, since, presumably, the average individual is aware that the upper hearing limit for a young adult is in the order of 17 kHz, and the amplifier should extend this high in frequency capability. In much scientific and engineering work, however, frequency is indicated graphically in angular frequency, $\omega = 2\pi f$, in radians per second.

8-8. AMPLIFIER BANDWIDTH

To indicate how constant an amplifier's gain is with frequency variation, we may specify the range of frequencies over which the gain does not deviate more than ±3 dB from the gain at the reference frequency. This is shown in Fig. 8-2, where the 3 dB bandwidth is given by Δf:

$$\Delta f = f_2 - f_1 \cdot \qquad (8\text{-}6)$$

For audio amplifiers this bandwidth should extend from 20 to 20,000 Hz, since this is the nominal frequency range to which the ear is sensitive. However, in video amplifiers and ac voltmeters, the bandwidth may have to extend from several hertz to tens of megahertz, which needs special compensating techniques. The 3 dB points, f_1 and f_2 in Fig. 8-2, are also referred to as half-power points, since a drop of 3 dB represents a power level of one-half

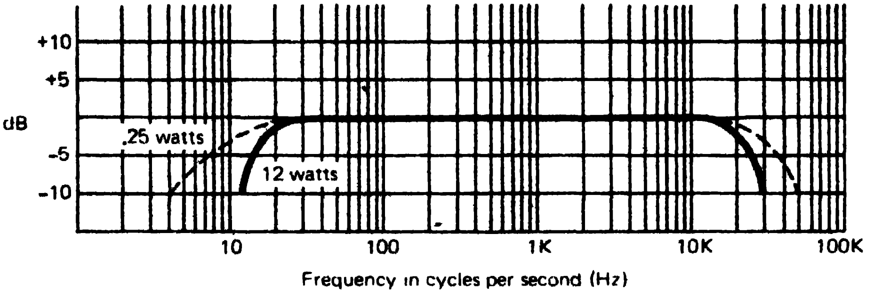


Fig. 8-3. Frequency response curve for a 12-watt audio amplifier.

that at the reference frequency. Although this may seem a large drop, the change is not easily noticeable to the average listener (a 1-dB change is barely perceptible), so that an amplifier may be considered to have a flat response from f_1 to f_2.

8-9. USE OF AN OUTPUT POWER METER

Another instrument that is useful in obtaining a frequency response curve of an amplifier is an output power meter. Basically, the instrument is a variable load resistor, (0.6 ohm to 30 K ohms), that may be connected to the output of an audio amplifier to show directly on a meter the power developed in the load, both in watts and decibels. The variation in power output (in either watts or decibels) with frequency may be obtained by monitoring the meter over the frequency range of the instrument, which is usually from 20 Hz to 20 kHz. The output power meter is also very useful in determining the load requirements for maximum power transfer, since the load may be varied until maximum indication of power on the meter is obtained. This provides a means of determining an unknown source impedance.

A simplified block diagram of an output power meter is shown in Fig. 8-4(a). The heart of the instrument is a variable-tap transformer that provides a different load, R_o, to the amplifier by changing the turns ratio between input and output, α, as shown by the equation, $R_o = \alpha^2 R_L$. Compensating resistors are added as the winding taps are changed to keep R_L constant. In this way, the square of the voltage developed across R_L is proportional to the power delivered by the amplifier,

$$\frac{V^2}{R_L} = P_o.$$

After attenuation, this voltage is rectified and filtered to provide a visual indication of power and decibels on a meter.

A Type 1840-A Output Power Meter is shown in Fig. 8-4(b). It is manufactured by General Radio Company and measures power in five ranges from 0.1 mW to 20 W. A decibel scale on the same meter reads from −15 to +43 dB, with 1 mW the reference level of 0 dB. The instrument's range may be increased to 200 W by inserting a T-network attenuator between the amplifier's output and the instrument's input, as shown in Fig. 8-5. The attenuator requires that 10 dB be added to the meter-switch indication on the instrument. The instrument may also be used to advantage in the measurement of insertion loss such as when a filter, tone control, or other device is added to an audio system. The difference in

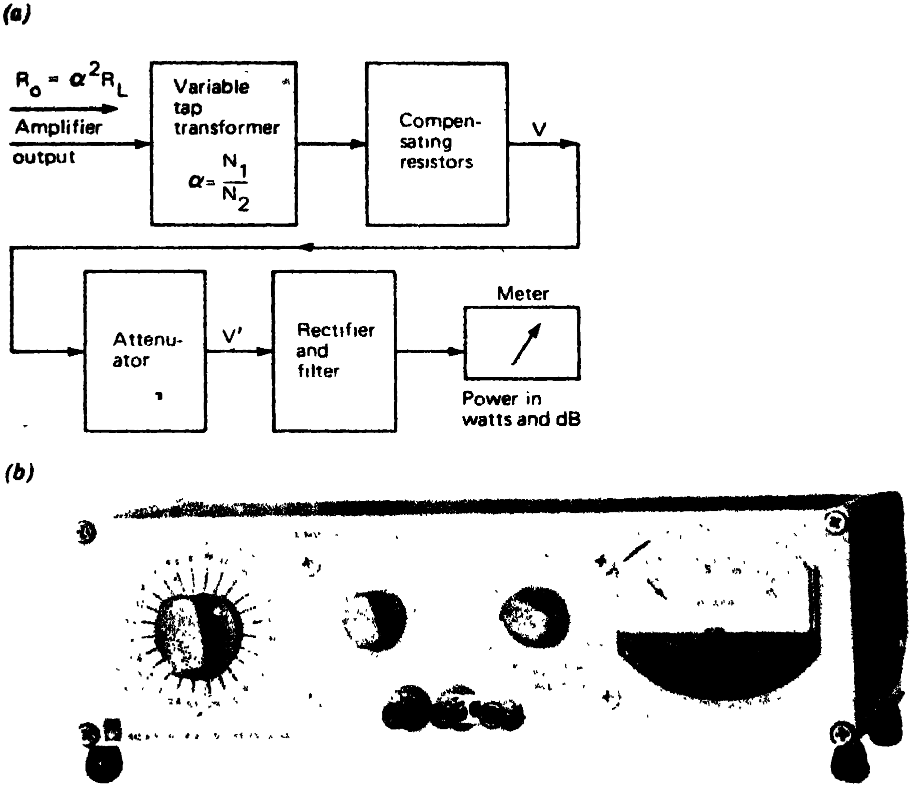

(a)
$R_o = \alpha^2 R_L$
Amplifier output
Variable tap transformer
$\alpha = \frac{N_1}{N_2}$
Compensating resistors
V
Attenuator
V'
Rectifier and filter
Meter
Power in watts and dB
(b)

Fig. 8-4. (a) Simplified block diagram of an output power meter.
(b) Commercial output power meter (Courtesy, General Radio).

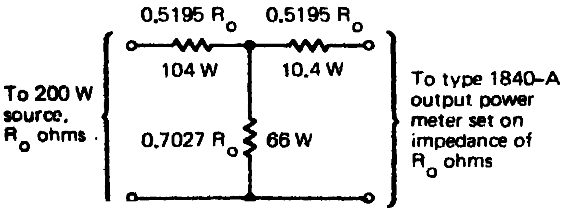

0.5195 R_o
0.5195 R_o
104 W
10.4 W
To 200 W source, R_o ohms
0.7027 R_o
66 W
To type 1840-A output power meter set on impedance of R_o ohms

Fig. 8-5. T-network attenuator to extend the range of the output power meter.

maximized decibel readings before and after inserting the device will yield the insertion loss in decibels.

8-10. PHASE RELATIONSHIPS IN AMPLIFIERS

Somewhere in the midfrequency range of an audio amplifier, the output is usually either in phase with the input waveform or 180° out of phase, as shown in Fig. 8-6. At other frequencies, however, the phase shift may be different and may vary with frequency as shown in Fig. 8-7. A review of the measurement of phase shift between two voltages, using the method of Lissajous figures on the oscilloscope, is covered in Appendix E.

Both the voltage gain, A_v, and the phase shift, ϕ, can be plotted simultaneously as a function of frequency on polar-coordinate paper. This consists of concentric circles with lines drawn from the center at 1° intervals, from 0° to 360°. The distance from the origin gives the gain magnitude (to scale), and the angle of the ray from the horizontal reference gives the phase shift angle. A sketch of such a plot is illustrated in Fig. 8-8.

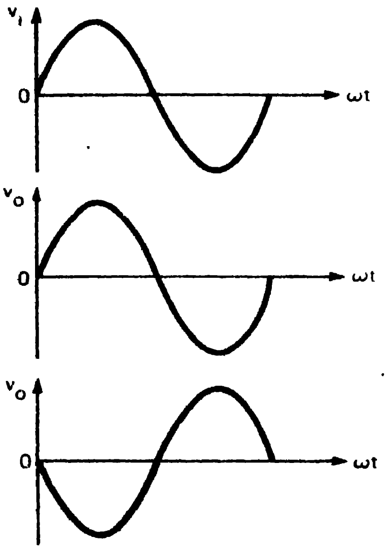


Fig. 8-6. (a) Input voltage waveform. (b) Output voltage waveform in-phase with input. (c) Output voltage waveform 180° out of phase with input.

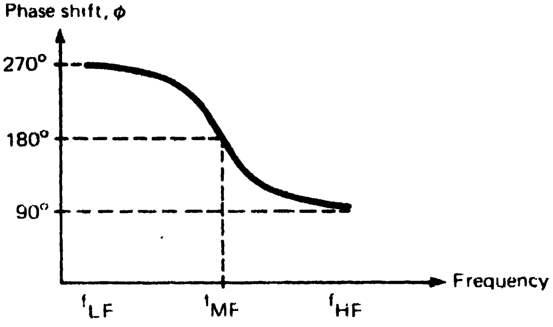


Fig. 8-7. Typical phase shift variation with frequency for an audio amplifier.

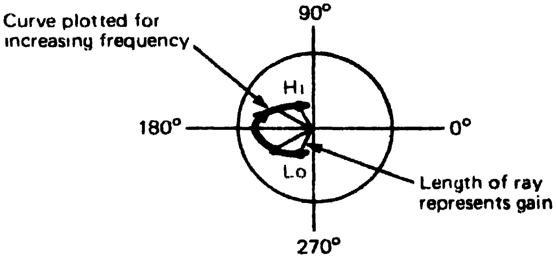


Fig. 8-8. Polar plot for an audio amplifier showing gain and phase.

A modified version of the polar plot in the complex frequency domain is called a Nyquist diagram and may be used to determine the stability of the amplifier (whether it will break into oscillations or not), but that is beyond the scope of this text.

8-11. SQUARE-WAVE TESTING

Instead of applying sinusoidal signals of different frequencies to an amplifier to observe its frequency response, a square wave may be used to show quickly the effects of both low and high frequency distortion. Since an audio amplifier is used to amplify complex waves containing many different frequencies at the same time, the amplifier should be tested under these conditions. For example, it is possible for an amplifier to adequately handle a 200 Hz signal or a 2,000 Hz signal separately. But when they are both present simultaneously it is possible for an interaction to occur between them, yielding an output different from the input. This is called intermodulation distortion, which was mentioned in Section 7-16.

It can be shown that a square wave is rich in odd harmonics, the equation giving

$$V_m \left(\sin \omega t + \tfrac{1}{3} \sin 3\,\omega t + \tfrac{1}{5} \sin \text{[illegible]}\,\omega t + \tfrac{1}{7} \sin 7\,\omega t + \ldots \text{ etc.}\right). \qquad (8\text{-}7)$$

Thus if a 1,000 Hz square wave is used, the amplifier is simultaneously being subjected to 3,000 Hz (one-third the amplitude of the fundamental), 5,000 Hz (one-fifth amplitude), 7,000 Hz, 9,000 Hz, etc., in addition to the fundamental frequency of 1,000 Hz.

Thus, if a square wave input is used, say at 10 Hz, then 1,000 Hz and finally at 10,000 Hz the shape of the output waveform will show the frequency response characteristics and phase distortion. Some of the output waveforms obtainable from an amplifier with separate bass and treble tone controls (see Section 8-14) are indicated in Fig. 8-9. It is to be noted that a tilt of the waveform indicates that some frequency components are experiencing an excessive lead or lag in phase while the squareness of the wave on the leading or trailing edge is an indication of high frequency reproduction capability. Linearity of low frequency response is indicated by the flatness of the output wave.

Although not usually found in amplifier testing, the sawtooth waveform in Fig. 8-10 is even richer in harmonics:

$$v = V_m \left(\sin \omega t + \tfrac{1}{2} \sin 2\,\omega t + \tfrac{1}{3} \sin 3\,\omega t + \tfrac{1}{4} \sin 4\,\omega t + \ldots \text{ etc.}\right). \qquad (8\text{-}8)$$

8-12. HEARING RESPONSE, LOUDNESS LEVELS

Before considering other general amplifier characteristics, let us look at the frequency response of the ear. The human ear responds to frequencies unequally, depending upon the loudness level, as shown in Fig. 8-11. The curves in Fig. 8-11 were determined by Fletcher and Munson,* who subjected a large number of listeners to a series of tests with pure tones. They were asked to match, in a noise-free chamber, tones of various frequencies with a standard 1000-Hz tone on the basis of equal loudness. The threshold of hearing was used as a loudness reference level, called zero phons, and corresponded to a pressure of 0.0002 dyne/cm^2 (or a power of 1[illegible] watts/cm^2) at the ear, for a 1000-Hz tone. (It has since become accepted as an average threshold of hearing by the American Standards Association, but initially represented the threshold of audibility for young ears.)

It can be seen that at the threshold of hearing, the ear is most sensitive to frequencies from 1 kHz to 6 kHz. For example, at the zero phon level it requires an increase in sound pressure level of 38 dB (0 dB = 0.0002 dyne/cm^2) for a 100 Hz tone to sound as loud as a 1 kHz tone. However, if the loudness

---

*Fletcher & Munson, *Speech and Hearing in Communication*, Van Nostrand, 1953.

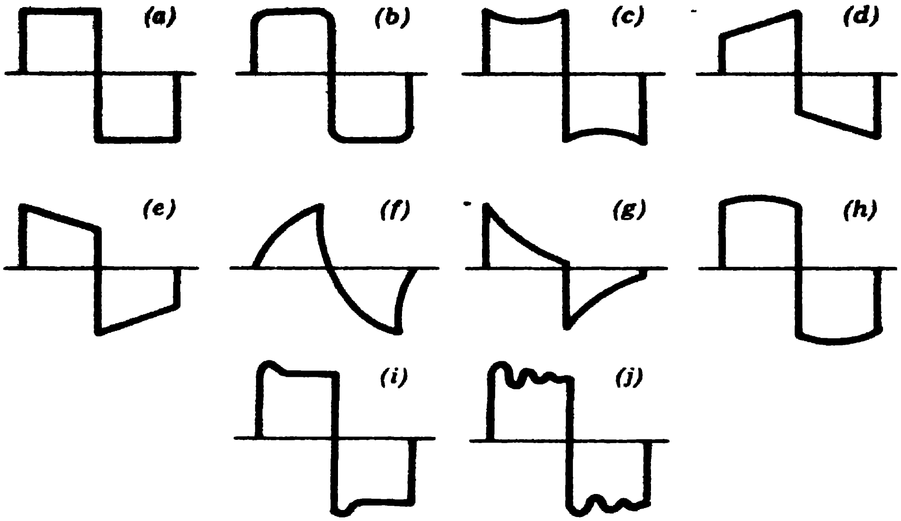


Fig. 8-9. Diagrams of amplifier response to square waves.

(a) Pure square wave input and output with no phase or frequency distortion.
(b) High frequency attenuation.
(c) Low frequency attenuation.
(d) Low frequency phase distortion
(e) High frequency phase distortion.
(f) High frequency attenuation and phase distortion.
(g) Low frequency attenuation and phase distortion.
(h) Excessive low frequency boost.
(i) Excessive high frequency boost.
(j) An indication of amplifier instability called ringing.

(Note: Not all of the above responses would be observed at the same frequency.)

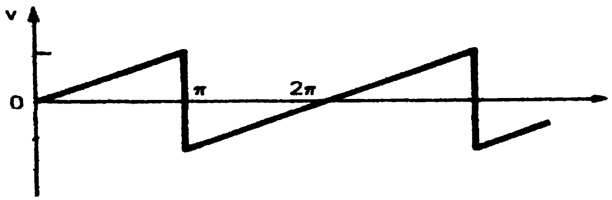


Fig. 8-10. Graph of a sawtooth waveform, which is rich in even and odd harmonics.

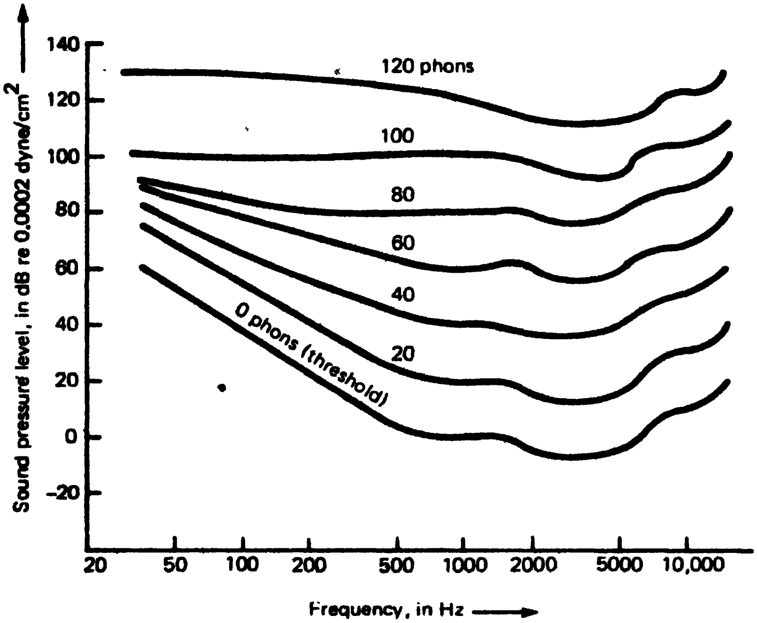


Fig. 8-11. Contours of equal loudness of pure tones. [From Speech and Hearing Communication *by H. Fletcher (c) 1953 by Litton Educational Publishing, Inc. (Published by Krieger Publishing.)]*

level is raised to 100 phons (corresponds to a 100 dB increase at 1 kHz) the ear is practically equally sensitive to all frequencies from 30 Hz to 16 kHz. A loudness level of 120 phons is termed the threshold of feeling, since at 20 dB higher the sensation is one of pain instead of hearing, and the auditory organs are susceptible to organic damage because of the extremely high intensity of air wave pressure.

Based on the reference level mentioned above for 0 dB, Table 8-2 indicates the noise or sound level for different locations and situations.

TABLE 8-2. OVERALL SOUND LEVELS IN DECIBELS (0 dB = 10^{-16} watt/cm^2)

LOCATION	SOUND LEVEL IN dB
F-84 at takeoff	135
Pneumatic riveter	125
Trumpet auto horn	110
Automatic lathe	100
10 hp outboard motor	90
Busy street	70
Ordinary conversation	65
Quiet auto	50
Quiet radio	40
Average whisper	20
Rustling leaves	10
Threshold of hearing	0

8-13. SOUND LEVEL METERS

These levels can be measured directly on a sound level meter, which consists of an omnidirectional microphone, a calibrated attenuator, an amplifier, standard weighting networks, and an indicating meter. The sound level range is from 24 to 150 dB (0 dB = 0.0002 μbar or 0.0002 dyne/cm^2). The weighting networks, A, B, and C, have frequency-gain characteristics that are the inverse of the 40-, 70-, and 120-phon curves respectively in Fig. 8-11, so that the meter provides an indication that corresponds approximately to the psychological loudness level of the sound. Fig. 8-12 shows a commercial sound level meter.

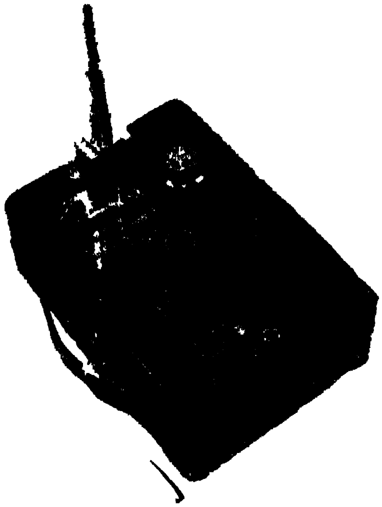


Fig. 8-12. A commercial sound level meter. (Courtesy of General Radio)

8-14. AMPLIFIER TONE CONTROLS

Thus far, we have found that an ideal amplifier should respond to all frequencies in the audio range with an equal gain. Thus, the amplifier will reproduce whatever material is fed into it without unnatural accentuation or reduction at any particular frequency, whether the material comes directly from a microphone or is intelligence from a phonograph, tape, or radio receiver. However, as we saw in Section 8-12, the ear is not equally sensitive to all frequencies, especially if the loudness level is quite low. As a result, a person who is listening to some program material with his amplifier at a low volume setting (in deference to his neighbors) will not be able to hear the bass and treble frequencies as loud as the midrange frequencies. Since these frequencies, particularly the bass frequencies, seem to add depth and fullness to music, it would be convenient if only these frequencies could be accentuated without increasing the overall volume of all frequencies. This is done by the use of separate bass and treble controls. (Of course this will add somewhat to the volume, but it is quite probable that the desired low and high frequency response will be achieved with less disturbance to the neighbors than if it were obtained by merely increasing the overall volume setting.)

Circuits providing tone control are of two types:

(1) Passive, involving only resistors, capacitors, and coils to attenuate frequencies at different levels in the frequency range.

(2) Active, or dynamic, using tubes or transistors to provide different gains for different frequencies, or by using feedback methods.

The passive circuits provide only attenuation of different frequencies. That is, no amplification is possible in a passive circuit. However, because of the attenuation caused by the passive circuits, amplifier stages are necessary to bring the signal level back up to that required to drive the subsequent power amplifiers.

A typical tone-control circuit with separate bass and treble controls is shown in Fig. 8-13. The bass and treble controls in the 0 position, shown in Fig. 8-13, establish a reference

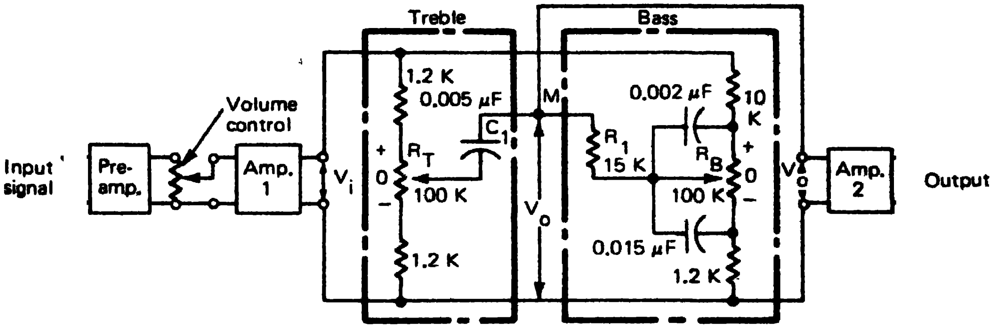

Treble
Bass
Input signal
Pre-amp.
Volume control
Amp. 1
V_i
1.2 K
0.005 μF
C_1
R_T
+
0
–
100 K
1.2 K
M
V_o
R_1
15 K
0.002 μF
10 K
R_B
100 K
0.015 μF
1.2 K
V_o
Amp. 2
Output

Fig. 8-13. Tone control circuit with separate bass and treble controls.

frequency response curve similar to that drawn in Fig. 8-3. This is called the flat response and the "boosting" and "cutting" that takes place at various frequencies is done with respect to this reference, as illustrated in a typical tone control frequency response curve in Fig. 8-14.

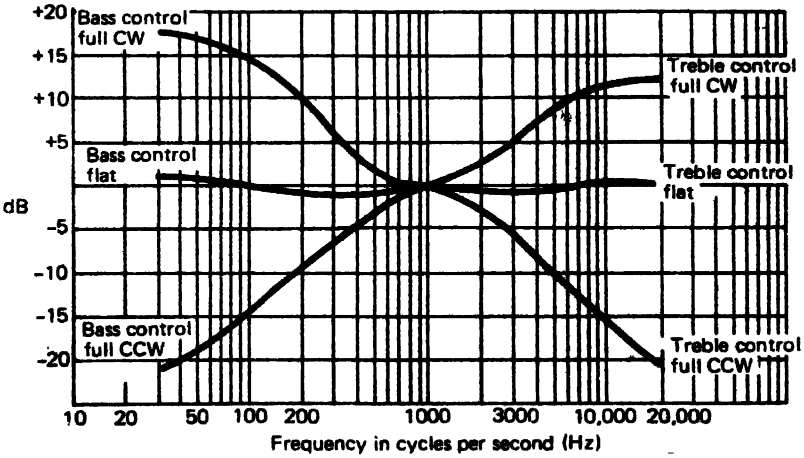


Fig. 8-14. Typical tone control frequency response curve for a 12-W amplifier.

The circuit of Fig. 8-13 will provide up to a 10-dB boost or a 15-dB cut for deep bass and high treble frequencies, depending upon the load. This is achieved in the following manner. The output from Amplifier 1 is fed simultaneously to the treble and bass controls. The action of C_1 is to prevent the passage of all but the higher frequencies, so that R_T provides treble control. If R_T is moved to (+), more of the input voltage, V_i, is developed across the voltage dividing circuit and appears at V_o. This provides a high frequency "boost." Similarly, the highs are attenuated or "cut" by moving R_T to (–).

The mid and low frequencies pass through the bass control and R_1 with the midrange frequencies passing through the 0.002 μF and 0.015 μF capacitors relatively unaffected. The low frequencies, however, are boosted by R_B moving to (+) and cut by a movement to (–), again because of a voltage dividing action. "M" represents the mixing point for the frequencies that pass through the bass and treble controls and constitutes the voltage V_o. In this circuit Amplifier 1 provides as much gain as is lost by the insertion of the bass and treble controls in their flat position.

8-15. RIAA EQUALIZATION

When the recording company cuts a record, they try to achieve a quieter recording by reducing the effects of surface noise. This they do by attenuating the low frequencies (to prevent overcutting of the grooves) and accentuating the high frequency tones (to give a better balance between high frequencies and surface noise). However, this intentional distortion (called the "recording characteristic") must be corrected by de-emphasis of the high frequencies and boosting of the lows at the playback amplifier. This compensation or equalization is made to conform to the standards of the RIAA (Record Industry Association of America, Inc.) which are shown in Fig. 8-15.

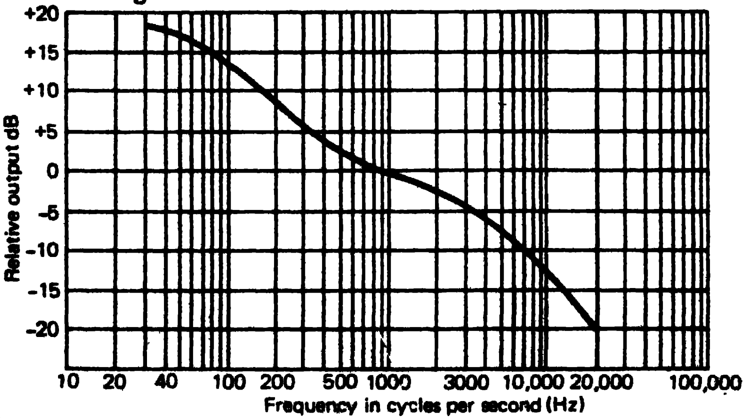


Fig. 8-15. RIAA equalization characteristic for discs.

The manner in which equalization is achieved depends upon the type of pickup used. A ceramic (or crystal) high-fidelity phonograph pickup is usually designed to provide the proper compensation when it is operated into the manufacturer's specified load resistance. Such a matching resistor is placed in series with the pickup and the preamplifier. A magnetic pickup, however, has essentially a flat frequency response characteristic, and a circuit to provide equalization is necessary. This may be at the input of the preamplifier (connected by the input selector switch) or incorporated in the amplifier's feedback network. A circuit which provides an emphasis of 18 dB for 50 Hz signals and a 12 dB de-emphasis of the 10,000 Hz signals is shown in Fig. 8-16.

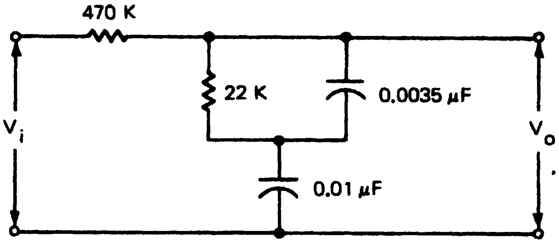


Fig. 8-16. A circuit to provide RIAA equalization for discs.

A similar "recording characteristic" is used in the production of magnetic tape recordings, and since the magnetic-tape pickup head also has a flat frequency response, an equalizing network and preamplifier is needed to provide equalization for the NARTB (National Association of Radio and Television Broadcasters) characteristic.

8-16. MISCELLANEOUS AMPLIFIER SPECIFICATIONS

Noise Figure

All resistors and active devices generate noise (see Appendix D, Section 16) which is amplified along with the input signal to the amplifier. If the ratio of signal to noise power at the output of an amplifier, $\frac{S_o}{N_o}$, is the same as the signal to noise power ratio at the input, $\frac{S_i}{N_i}$, the amplifier is said to be "noiseless." In general, the signal will be impaired due to the noise of the amplifier, and a measure of the degradation is called the noise figure, expressed in dB as

$$NF = 10 \log \frac{\frac{S_i}{N_i}}{\frac{S_o}{N_o}} . \qquad (8\text{-}9)$$

Thus, a noiseless amplifier would have a noise figure of unity or zero dB. Alternatively the hum and noise introduced by an amplifier may be expressed simply by

$$\text{Signal to Noise} = 10 \log \frac{S_o}{N_o} . \qquad (8\text{-}10)$$

For example, a 12-watt amplifier may have hum and noise 47 dB below 12 watts for a magnetic phonograph input, but 63 dB below 12 watts for the crystal and tuner inputs.

Dynamic Range

An amplifier's capability to accommodate the loudest and softest portions of recorded material without excessive distortion or noise is indicated by its "dynamic range." The dynamic range is defined as the ratio of the maximum usable output signal (for a distortion not in excess of 10%) to the minimum usable output signal (for a signal to noise ratio of at least 20 dB).

This is normally expressed in dB with 40 dB being acceptable for an amplifier, but 70 dB an exceptionally good figure.

Input Sensitivity

The input voltage requirement to produce a given output power is called the sensitivity of the amplifier and is shown in Fig. 8-17 for various types of inputs for a 12 watt amplifier.

Power Rating

Audio amplifiers may have their power rating expressed in rms watts; this represents the power delivering capability when handling a a continuous sine wave. However, since music and speech seldom consist of sustained sine waves, but rather complex waves of lower rms value, the amplifiers may be rated by the EIA or the IHFM (Institute of High Fidelity Manufacturers) in watts of music power that is higher than the rms power. For example, an amplifier rated at 25 watts of rms power was given a music power rating of 45 watts by the EIA, while the IHFM rated the same amplifier at 50 watts music power.

Another popularly advertised rating indicates the peak power output. This is the power at the peak of the output signal, and since

$$P = \frac{V^2}{R} \text{ then } P_{peak} = \frac{(V_{peak})^2}{R}$$

$$= \frac{(\sqrt{2}\,V)^2}{R} = \frac{2\,V^2}{R} = 2\,P.$$

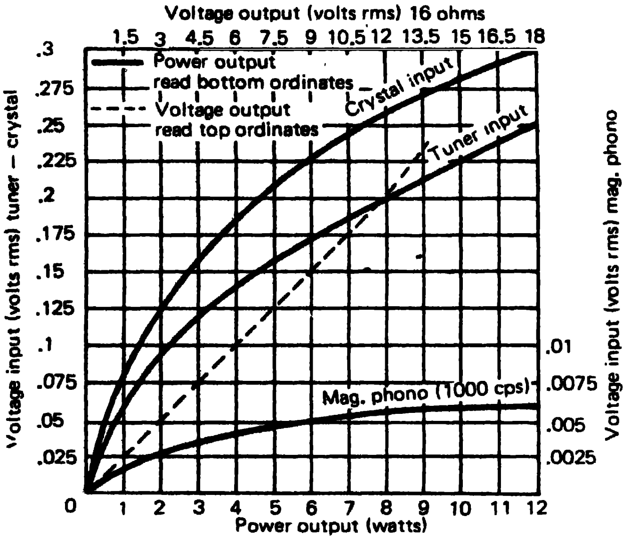


Fig. 8-17. Sensitivity for various inputs of a 12 W amplifier.

Thus this amplifier could claim to have a peak music power rating of 100 watts. If this were a stereo amplifier with two of the above amplifiers the preceding figures would be doubled, and referred to as total power.

8-17. LOUDSPEAKER REQUIREMENTS

Finally, a word on the tendency to large power amplifiers in home entertainment systems. The *acoustic* power from a loudspeaker necessary to reproduce the loud portion of orchestral music is in the order of 0.5 watt in the average home. However, loudspeakers with their multiple transducers of woofer, tweeter and associated crossover circuits seldom exceed 5% in efficiency — perhaps as low as 2% —, so the input power to the speaker must be approximately $\frac{0.5}{0.02}$ W = 25 W.

When an amplifier has this much gain capability, it is important to make sure that the noise level in the amplifier is not excessive at the required listening level. Most phonograph records and program material from broadcast radio stations have a difference between the loudest and softest sounds of approximately 60 dB. Thus, the noise of the amplifier itself (hum, transistor noise, etc.) must be at least 60 dB below the signal level at the required listening level, so that when the quiet passages are played the noise will not be audible. An amplifier with a very large power capability and dynamic range playing at a fraction of its rated value will be able to handle very easily a wide range in volume without "straining," keeping the hum and noise completely inaudible.

8-18. SUMMARY

(1) Decibels are used to provide a comparison of two power levels, and in an amplifier the gain may be obtained by any of the following equations:

$$\text{dB} = 10 \log \frac{P_o}{P_i},$$

$$\mathrm{dB} = 20 \log \frac{V_o}{V_i} + 10 \log \frac{R_i}{R_o},$$

$$\mathrm{dB} = 20 \log \frac{I_o}{I_i} + 10 \log \frac{R_o}{R_i}.$$

(2) In stating an absolute power level in decibels, it is necessary to define the zero decibel reference level, which in much electronic work is taken to be 1 mW of power in a 600-ohm resistance load.

(3) When using a voltmeter to determine an absolute power level in decibels, it is necessary to add a voltmeter range correction factor and an impedance correction factor to the direct reading from the voltmeter.

(4) A frequency response curve of an amplifier is a graph showing how the output varies with frequency, with the output at 1 kHz usually taken to be the reference of 0 dB.

(5) An audio amplifier's bandwidth normally extends from 20 to 20 kHz with a variation of not more than ±1 dB (±3 dB at the most) from the reference.

(6) The phase shift between input and output of an audio amplifier varies with frequency and may be drawn graphically in conjunction with amplifier gain on polar coordinates.

(7) A square wave is rich in odd harmonics and may be used to quickly show the frequency response of an amplifier by observing the output for squareness.

(8) The ear responds differently to different frequencies, depending upon the loudness level; consequently bass and treble tone controls are often used to boost an amplifier's response at low and high frequencies.

(9) A "recording characteristic" is the intentional frequency distortion made during recording to reduce the effects of surface noise, and must be compensated for in the playback amplifier by equalizing networks to conform to RIAA or NARTB standards.

(10) To drive large, low efficiency loudspeakers, amplifiers with a large power rating are necessary; such amplifiers must have a dynamic range to accommodate the loudest and softest material without excessive distortion or noise.

QUESTIONS

Q8-1. Was the decibel defined in terms of (a) a voltage level or (b) a power ratio?

Q8-2. What power ratio is one bel?

Q8-3. If the electrical power to a loudspeaker is increased from 2 watts to 4 watts, what effect would your ear notice?

Q8-4. Under what conditions is the power gain in decibels equal to $20 \log \left(\frac{V_o}{V_i}\right)$?

Q8-5. A leading manufacturer of voltmeters includes the following statement pertaining to the use of the decibel scale on a voltmeter for measuring *gain*: "If the signal voltage at the input of an amplifier read 0 dB in the 1.5 volt position and the output voltage read 0 dB in the 15 volt position, it would indicate that the amplifier has a *gain* of 20 dB." Comment on this statement, with reference to Q8-4, clarifying what is meant by *gain*.

Q8-6. If a network is said to have a power gain from input to output of -3 dB, what does this mean?

Q8-7. Why is it not possible to use 0 watts as the zero decibel reference level?

Q8-8. What correction factors are necessary to obtain the absolute decibel level above 0 dBm from a voltmeter reading?

Q8-9. Describe in your own words how you would use a voltmeter with a decibel scale to obtain the frequency response curve of an audio amplifier.

Q8-10. What is the -3 dB bandwidth of the amplifier whose frequency response curve is shown in Fig. 8-3?

Q8-11. What is the advantage of using semi-log paper for a frequency response curve as opposed to a linear scale?

Q8-12. How may an output power meter be used to determine an audio amplifier's output resistance?

Q8-13. From an examination of the waveforms in Fig. 8-9, what frequency components are responsible for:

(a) square corners?
(b) flat tops?

Q8-14. What range of frequencies is the human ear most sensitive to (± 3 dB) at:

(a) the threshold of hearing, 0 phons?
(b) a loudness level of 100 phons?

Q8-15. How would a sound level meter indication on the C weighting range compare with the reading on the A weighting (same location and source)?

Q8-16. Explain in your own words why bass and treble controls are useful additions to an audio amplifier.

Q8-17. If the output of Amplifier 1 in Fig. 8-13 were connected directly to the input of Amplifier 2, how would the output of Amplifier 2 compare with the output with the tone control circuit in place? Explain.

Q8-18. Why is it necessary to provide an equalization or de-emphasis circuit in an amplifier that is playing back intelligence from a disc or tape using a magnetic pickup?

Q8-19. What does a high dynamic range indicate in an amplifier? Does the dynamic range involve the amplifier's noise figure?

Q8-20. List the different power ratings that a manufacturer could advertise for an amplifier. Which do you think is the most honest to use in comparing amplifiers?

PROBLEMS

P8-1. What power ratio is (a) one bel? (b) one decibel?

P8-2. A 5 mV input to the magnetic phono input of an amplifier which has an input resistance of 47 K ohms produces a power output of 6 watts. Calculate the power gain of the amplifier.

P8-3. An amplifier is known to have a power gain of 40 dB. If the output is 4 watts determine the input power.

P8-4. What output power is obtained from an amplifier whose power gain is 55 dB when the input power is 1 mW?

P8-5. An audio amplifier has an input impedance of 160 K ohms. When a load of 8 ohms is connected across the output of the amplifier a 0.5 V (rms) input signal causes a 15 V range voltmeter to deflect to 5 V (rms).

(a) Calculate the power gain of the amplifier in dB.
(b) If the voltmeter has a dB scale calibrated for 0 dB = 1 mW in 600 ohms, as in Fig. 8-1, determine the output power level of the amplifier in decibels, with reference to the RIA 0 dB level. (Hint: use *both* correction factors along with the direct reading from the dB scale.)
(c) If the input voltage is increased to 0.5 $\sqrt{2}$ volts from 0.5 V what will be the new indication of the voltmeter on the dB scale (direct reading) and the power gain of the amplifier?

P8-6. A filter is inserted into a communications line. The input to the filter is 600 ohms and the output is 750 ohms. If an input voltage of 0.5 V causes an output of 0.35 V, calculate the dB loss across the filter.

P8-7. An amplifier that has a power gain of 36 dB provides an output of 10 V peak-to-peak when a 16 V peak-to-peak input is applied. Determine the ratio of input to output resistance.

P8-8. A voltmeter across the input to an amplifier of input resistance 47 K ohms indicates −3 dB on the 5 V range. When connected across the output load of 1 K ohm, the voltmeter on the 15 V range indicates +5 dB. What is the dB power gain of the amplifier?

P8-9. The power output of a given amplifier is 1 watt. What is the relative level of this power in decibels using as the 0 dB reference

(a) the EIA 0 dB reference of 6 mW in 500 ohms?
(b) the RIA 0 dB reference of 1 mW in 600 ohms?

P8-10. Determine the ±3 dB bandwidth of the amplifier whose frequency response curve is shown in Fig. 8-3 for the following conditions:

(a) 0.25 watts output.
(b) 12 watts output.

P8-11. If the amplifier whose response is given by Fig. 8-3 is delivering 0.25 watts at 1 kHz, what will be the power in watts at a frequency of:

(a) 7 Hz? (b) 40 Hz? (c) 30 kHz?

P8-12. If the amplifier whose response is given by Fig. 8-3 is delivering 12 watts at 1 kHz, what will be the power in watts at a frequency of:

(a) 15 Hz? (b) 40 Hz? (c) 30 kHz?

P8-13. An output power meter connected to an amplifier is adjusted until the output power indication is a maximum of 10 watts at a load resistance of 8.1 ohms. Determine:

(a) the decibel indication if 0 dB = 1 mW.
(b) the open circuit voltage and output resistance of the amplifier.
(c) the decibel indication on the meter when the load resistance is reduced to 6 ohms.

P8-14. The T-network attenuator shown in Fig. 8-5 is connected between a source and the output power meter set on an impedance of 10 ohms. Determine:

(a) the input resistance seen by the source.
(b) the power input to the power meter if the source supplies 200 W.

P8-15. *Sketch* two cycles of the first three terms in Eq. 8-7 to scale, using $V_m = 10$ V. Add them algebraically to approximate a square wave. Can you see how it will take many odd harmonics to give a *good* square wave?

P8-16. A ram jet has an acoustic power in the region of 100,000 watts while the voice, in a very soft whisper, is in the order of 1×10^{-9} watts.

(a) Determine the sound level in decibels that the ram jet is above a soft whisper.
(b) If the soft whisper is considered to be 20 dB above the threshold of hearing, determine the acoustic power at 0 dB.

P8-17. With reference to the curves in Fig. 8-14, assume the tone controls are flat and the amplifier has an output of 0.5 W for any single input frequency from 100 Hz to 10,000 Hz. If the bass and treble controls are both turned fully clockwise, what would be the output power for:

(a) a single input frequency of 100 Hz?
(b) a single input frequency of 10 kHz?
(c) a single input frequency of 1 kHz?

P8-18. Using circuit analysis techniques, determine the ratio of V_o to V_i in Fig. 8-16 for the following frequencies:

(a) 1 kHz. (b) 100 Hz. (c) 10 kHz.

Determine the relative output in dB for (b) and (c) compared with 1 kHz (0 dB). Compare with the values shown in Fig. 8-15 for the same frequencies.

P8-19. (a) If a 12-W amplifier has a noise level 47 dB below 12 W for a magnetic-phono input and 63 dB below 12 W for a tuner input, determine the noise at the output in watts when 12 W are being delivered for each of the two cases.

(b) What is the signal to noise power ratio in both cases?

(c) If the signal to noise power ratio at the magnetic input is 10^7, determine the noise figure, NF, of the amplifier for this condition.

P8-20. Determine the dynamic range of an amplifier in decibels given the following:

(a) The amplifier has a sensitivity given by Fig. 8-17 and can accept 0.225 volts at the tuner input without excessive distortion at the output (less than 10%).

(b) The amplifier's signal to noise ratio gets progressively smaller as the input signal is reduced, reaching a value of 20 dB when the output is 10 mW.

PART 3

BIPOLAR TRANSISTOR AMPLIFIERS

9

BASIC CHARACTERISTICS OF THE TRANSISTOR

9-0. INTRODUCTION

Transistor theory depends heavily upon the semiconductor theory covered earlier under PN junctions, since a transistor consists essentially of two PN junction diodes back-to-back, or front-to-front. It is from this standpoint that the following discussion will proceed, to obtain a physical feeling of transistor operation instead of getting lost in equations that describe hole and electron density, etc. However, later work will cover equivalent models of the transistor which will help predict the operation of the transistor in an amplifier.

9-1. BASIC TRANSISTOR AMPLIFIER

For those who have studied tubes, a comparison between a transistor and a triode may be helpful, although this discussion should also provide a useful introduction for those unfamiliar with tubes.

Although the transistor is different from the vacuum triode in that it is a current-operated device instead of a voltage controlled device, nevertheless, a likeness can be found between the two. Refer to the circuits of Fig. 9-1. In this case the transistor consists of a P-type material sandwiched between two N-type materials, and is thus called an NPN type.

The symbol for the transistor shows an arrow on the emitter which points in the direction of conventional current flow or holes in the transistor. It can be seen that the emitter can be compared with the cathode (although no heat is required), the collector to the plate since it collects electrons from the emitter, and the base to the control grid, since, depending upon the amount of current injected into the base, it controls the current flowing in the output or collector circuit.

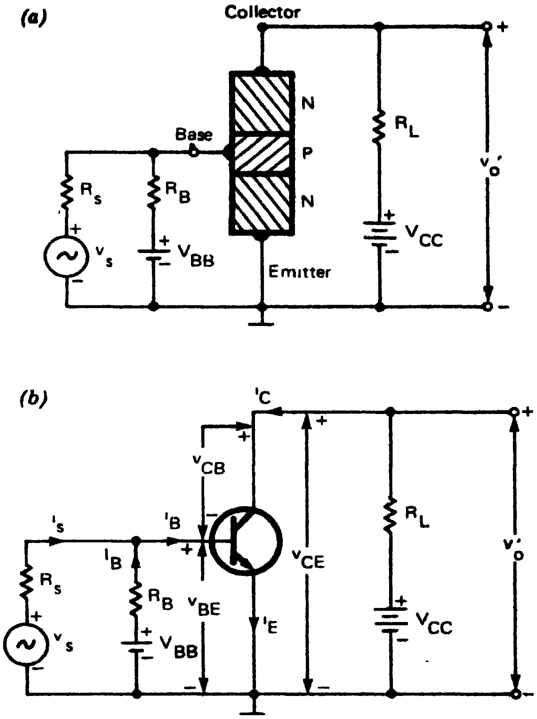


Fig. 9-1. (a) Basic common emitter amplifier showing NPN transistor in diagrammatic form. (b) Amplifier showing transistor symbol and voltages.

The particular circuit shown is a common- (or grounded-) emitter (CE), and is the counterpart of the grounded-cathode vacuum tube amplifier. As in the vacuum tube amplifier, a load resistor R_L is required to develop the output voltage due to the collector current i_C flowing through it. V_{CC} is the supply voltage, typically 12 to 20 V in many amplifiers. The voltage from collector to emitter is referred to as v_{CE}, and the voltage from base to emitter as v_{BE}. The double subscript notation is very common in transistor work, and the small letter for v designates instantaneous values. However, where it is clear that all voltages are with respect to the common line or ground (emitter in this case), the single subscript notation is sufficient and would be v_C and v_B here.

V_{BB} is used to bias the transistor in the center of its characteristics, much like a triode; it could have been placed in series with v_s just as it was in the triode amplifier. With no signal v_s applied, V_{BB} will supply a dc current into the base, I_B. Some corresponding dc value of collector current will flow, called I_C, the sum of the two giving the emitter current I_E.

Thus $I_E = I_B + I_C$ with no signal applied. The application of v_s will cause a current i_s to be added to I_B resulting in an ac component of current riding on a dc level, called i_B. This is depicted in Fig. 9-2. This in turn will cause an ac component of current i_c to flow in the collector circuit, which added to I_C is called i_C. The doping of the junctions and the width of the base are chosen so that a small base current i_B can control a larger collector current i_C. Hence, the transistor has the capability of current gain, A_i. Also, since the voltage from base to emitter is small and the voltage in the collector circuit is larger, then a voltage gain A_v is also possible. This gives the transistor a power gain $A_p = A_v A_i$.

It should be noted that as far as the transistor is concerned

$$i_E = i_C + i_B \text{ and } v_{CE} = v_{CB} + v_{BE}. \tag{9-1}$$

9-2. TWO-DIODE ANALOGY FOR A TRANSISTOR

It can be seen from the physical drawing of the transistor that there are two PN junctions, so that the transistor can be thought of as two diodes — in this case back-to-back, as shown in Fig. 9-3.

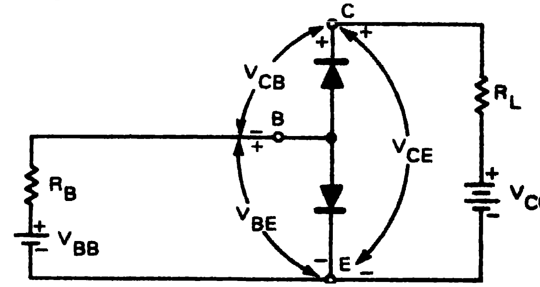


Fig. 9-3. Two-diode analogy diagram for an NPN transistor with bias and supply voltages.

Note that the diodes point outward as does the arrow on the emitter lead of this NPN transistor. Observe that the base-emitter diode junction is *forward*-biased by V_{BB}, but the collector-base diode junction is *reverse*-biased by V_{CC}. This is *always* the case for normal amplifier operation of a transistor in any of the three transistor amplifier configurations.

Note that, since $V_{CE} = V_{CB} + V_{BE}$ and that V_{BE} is very small (approximately 0.5 V), then $V_{CB} \simeq V_{CE}$.

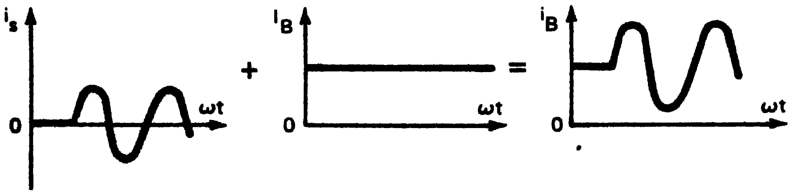


Fig. 9-2. Showing how the total base current i_B consists of an ac component i_s and a dc component I_B.

This diagram should not be taken to represent accurately a model for a transistor. It is used to emphasize the correct biasing of the transistor and will be used to help explain its operation later. However, resistors and a current generator could be added to this diagram to give what is called a piece-wise-linear T-equivalent model.

9-3. TRANSISTOR INPUT CHARACTERISTICS

Using the two-diode analogy, the input and output characteristic curves for the transistor can be explained. Since the base-emitter diode is forward-biased, it should be expected that V_{BE} will be small and in the region of 0.6 volt for a silicon transistor and 0.2 volt for germanium. How much current will flow in the base when V_{BE} is applied will depend upon the transistor type (signal or power). From the curve in Fig. 9-4, it can be seen that what happens at the input is dependent upon what voltage is present at the output, V_{CE}. This is what makes the transistor so difficult to analyze: the interactions between input and output. This difference in input characteristic when V_{CE} is applied is due to the change in the width of the depletion region under conditions of reverse bias.

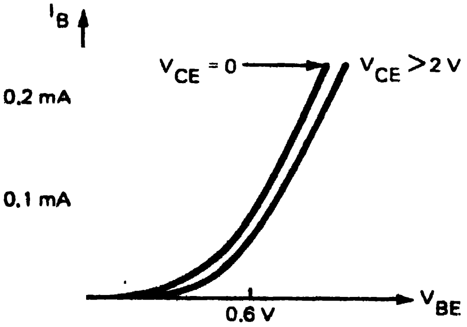


Fig. 9-4. Typical input characteristic for a silicon transistor showing how the base current vs. base-emitter voltage curve depends upon the collector or output voltage.

9-4. TRANSISTOR COLLECTOR CHARACTERISTICS, CE

In the output circuit, it may at first be asked how can *any* current flow in the collector circuit if the collector-base diode is *reverse* biased? The answer lies in the heavy doping of the collector with respect to the base and the narrow width of the base. This provides the *effect* of a "leaky diode" from collector to emitter, so that a large reverse current can flow *if* the base-emitter junction is forward biased. (This will be explained more adequately later.) However — and this is where the control in the transistor is found — , the amount of current that *will* flow in the collector circuit depends upon how well the base-emitter diode has been *turned on*. That is, the PN junction from base to emitter acts like a valve or control gate. The larger the forward bias V_{BE} and the base current I_B, the larger will be I_C for a given V_{CC}.

Thus, the output characteristic curves can be drawn, of I_C against V_{CE} with I_B as the control parameter, as shown in Fig. 9-5.

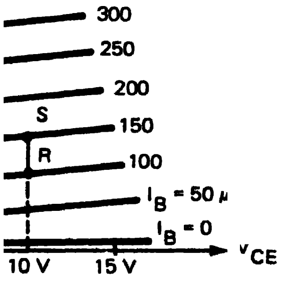


Fig. 9-5. Typical collector characteristic curves for a signal type of transistor in the CE connection.

These characteristic curves can be likened to the plate characteristics of a tube except that base *current* is the controlling factor instead of *grid voltage*. Note that the curves are more like a pentode's (slightly more upward sloping) than like a triode's.

9-5. COLLECTOR CUTOFF CURRENT, I_{CEO}

If no forward bias is applied ($I_B = 0$), then the only current that will flow in the collector circuit will be the leakage current that can flow through the base-emitter junction with the base open. This is referred to as I_{CEO}, the leakage or cutoff current from *c*ollector to *e*mitter with the base *o*pen. (Note the subscript letters!) See Fig. 9-6.

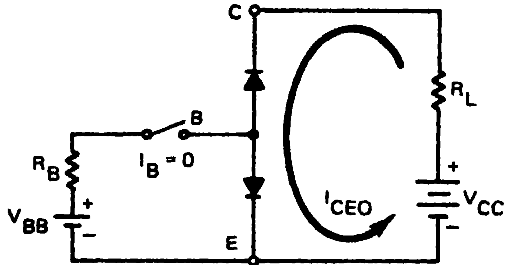


Fig. 9-6. Using the two-diode analogy circuit to explain the collector cutoff current, I_{CEO}.

9-6. FORWARD CURRENT TRANSFER RATIO, CE

The magnitude of the base current compared with the collector current in Fig. 9-5 suggests that the device has a good current gain in the common-emitter connection. This gain may be defined in terms of dc current or small signal ac current as follows:

Dc Current Gain

The dc current gain or the static forward-current transfer ratio in a CE is the ratio $\frac{I_C}{I_B}$ and may be referred to as β_{dc} or h_{FE}. That is:

$$\beta_{dc} = h_{FE} = \frac{I_C}{I_B}. \ (\beta = \text{beta.}) \tag{9-2}$$

A distinction is sometimes made between the large-signal current gain of a transistor, designated by β, and the dc current gain, β_{dc}. However, where the cutoff current I_{CBO} (Section 9-11), equal to

$$\frac{I_{CEO}}{\beta + 1},$$

is small compared with I_B and I_C, the values of β and β_{dc} are approximately equal.

This parameter, β, is used extensively in the chapters on thermal stability and is taken to be synonymous with h_{FE}. Thus:

$$\beta \approx \beta_{dc} = h_{FE} = \frac{I_C}{I_B}. \tag{9-3}$$

Evaluating β at point S of the characteristic curves in Fig. 9-5 yields the following:

$$\beta = \frac{I_C}{I_B} = \frac{7.5\ \text{mA}}{150\ \mu\text{A}} = \frac{7.5 \times 10^{-3}}{150 \times 10^{-6}} = 50.$$

Values for β (or h_{FE}, which is the designation used in many transistor manuals and catalogs) may vary from 10 to over 500 according to the transistor type. But even in transistors of the same type it is not unusual to experience a 3:1 spread between the minimum and maximum values of β.

Ac Current Gain

The small-signal gain or forward-current transfer ratio in a CE is the ratio $\frac{\Delta I_C}{\Delta I_B}$ and may be designated as β_{ac} or h_{fe}. That is:

$$\beta_{ac} = h_{fe} = \frac{\Delta I_C}{\Delta I_B}\ (V_{CE}\ \text{constant}), \tag{9-4}$$

where ΔI_C is the change in collector current resulting from a small change in base current ΔI_B at a fixed value of collector-to-emitter voltage, V_{CE}.

Evaluating h_{fe} in Fig. 9-5 between points R and S, a change in I_B from 100 μA to 150 μA causes a change in I_C from 5 mA to 7.5 mA, while V_{CE} is held constant at 10 volts. Thus

$$\beta_{ac} = h_{fe} = \frac{\Delta I_C}{\Delta I_B} = \frac{(7.5 - 5)\,\text{mA}}{(150 - 100)\,\mu\text{A}}$$

$$= \frac{2.5 \times 10^{-3}}{50 \times 10^{-6}} = 50.$$

h_{fe} is a parameter used in the analysis of small-signal amplifiers in Chapter 14 and may have a range of values similar to that mentioned above for h_{FE}.

A good amplifying transistor should have a large value for h_{fe} (corresponds to a large μ in a tube). Where the base current is in the order of milliamps or even amps as in a power transistor, the value of h_{fe} will normally be smaller, but in small-signal transistors h_{fe} may be as large as 500.

Relation Between Dc and Ac Current Gains

In the above examples it was observed that h_{FE} had a value numerically equal to h_{fe}. This is often the case and can be shown to be true over quite a large region of the characteristic curves by referring to Fig. 9-7. This is a typical graph showing how a silicon transistor's static forward current transfer ratio h_{FE}, or β, varies with collector current and how it also depends upon temperature. (Values of h_{FE} have been normalized with respect to the value at 25°C and a current of 10 mA.) h_{fe} is also frequency-dependent, but this will be covered in Chapter 16.

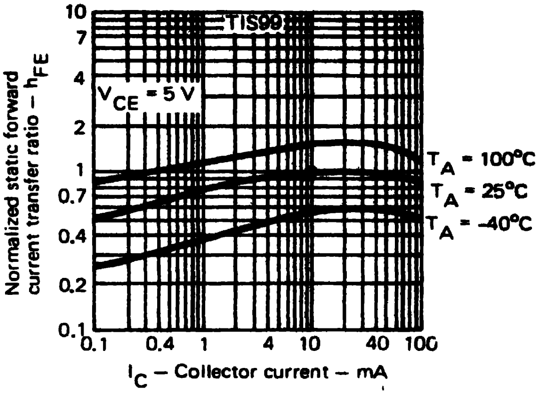


Fig. 9-7. Graph showing how β varies with collector current and temperature for a TIS99 silicon transistor. (Courtesy of Texas Instruments, Incorporated.)

Referring to the curves in Fig. 9-7, it can be seen that β has a maximum value in the region of I_C = 20 mA. Now it can be shown that to the left of this peak $h_{fe} > h_{FE}$, while for currents to the right of the maximum $h_{fe} < h_{FE}$. Thus, for a current range of I_C from 5 mA to 70 mA, there is less than a ±5% difference between h_{fe} and h_{FE} for this transistor at 25° C. And since the value of h_{fe} or h_{FE} may have some wide variation even for a given transistor type, we can often approximate and refer to a transistor's current gain as β and not make any distinction between ac and dc. Only at very low or very high values of collector current would any significant difference be apparent.

Using the definition of $\beta = h_{FE} = \dfrac{I_C}{I_B}$, we can write an equation for the collector current as a function of the base current:

$$I_C = \beta I_B. \tag{9-5}$$

And if the cutoff current is included:

$$I_C = \beta I_B + I_{CEO}, \tag{9-6}$$

which approximates the collector characteristics drawn in Fig. 9-5. Why I_C is dependent upon V_{CE} will be explained in Section 9-14.

9-7. PERMISSIBLE OPERATING AREA OF A TRANSISTOR, CE

It should be recalled that if a diode is reverse-biased there is a limit on the voltage that can be applied which is the avalanche voltage. When this voltage is reached in the collector-base diode, the collector current increases sharply, as indicated in Fig. 9-8.

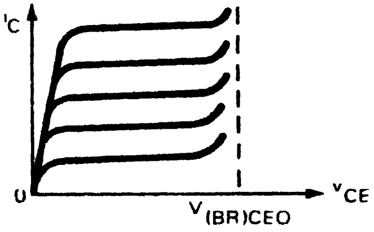


Fig. 9-8. CE characteristic curves showing the collector-to-emitter breakdown voltage, $V_{(BR)CEO}$.

The transistor manual gives the $V_{(BR)CBO}$, the *b*reakdown voltage from *c*ollector to *b*ase with the emitter *o*pen, as the value when V_{CB} reaches the avalanche voltage. This is typically 110 volts. However, the $V_{(BR)CEO}$, the breakdown voltage from *c*ollector to *e*mitter with the base *o*pen, is generally 40 to 50 percent of $V_{(BR)CBO}$, typically 40 volts in a 2N1611. This is the upper limit of V_{CE} that can be placed across the transistor without harming it.

There is also a limit to the collector current that the transistor can handle, which, with the maximum power dissipation capability, describes the permissible area of operation of the transistor. This is shown in Fig. 9-9. The maximum dissipation hyperbola is drawn through a number of points obtained by assuming a collector current and determining the maximum permissible voltage from collector to emitter without exceeding the power rating. Power ratings of transistors depend upon operating temperatures and vary from as low as 50 mW up to the hundreds of watts. The effects of temperature will be covered more thoroughly under thermal stabilization and under power amplifiers.

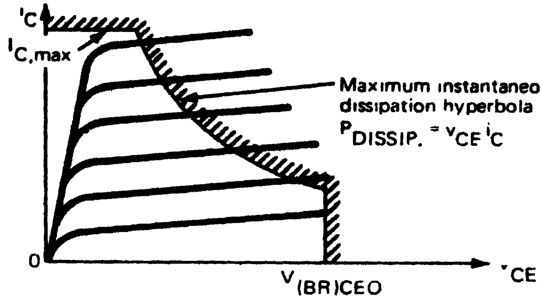


Fig. 9-9. Showing the permissible operating area of a transistor's CE characteristics.

9-8. THE BASIC COMMON-BASE AMPLIFIER, CB

It will be shown that the common-emitter amplifier yields the largest power gain of all the transistor amplifier configurations. However, just as the grounded-grid vacuum tube amplifier was suitable at high frequencies, so the grounded- or common-base transistor amplifier excels at high frequencies. The basic amplifier is shown in Fig. 9-10.

It should be noted that a PNP transistor has been used in this circuit (for variety), but the emitter-base PN junction is still forward-biased by V_{EE}, and the collector-base diode is reverse-biased by V_{CC}.

Without any signal applied, the emitter current is the dc bias current I_E. Most of this current is swept into the collector portion of the transistor, giving a collector current I_C and leaving only a small base current I_B.

Thus $I_E = I_B + I_C$ as before, and when v_s is applied this becomes $i_E = i_B + i_C$.

The same diode analogy can be made, but this time, since the arrow on the emitter points inward, the two diodes will also point inward, as illustrated in Fig. 9-11.

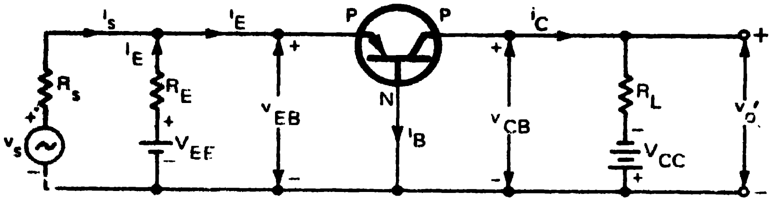


Fig. 9-10. Basic common-base amplifier with PNP transistor.

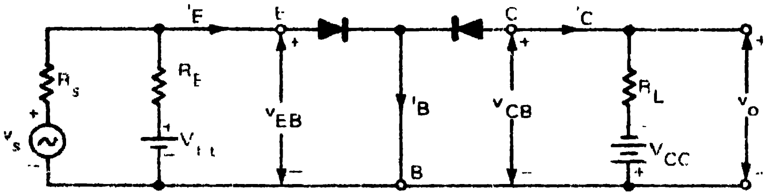


Fig. 9-11. Using the two-diode analogy to represent a CB amplifier

The input and output characteristics can again be drawn as in Fig. 9-12.

The usual forward-biased diode characteristic is found for the input, although it is seen that the input or controlling current is much larger than in the CE amplifier. The output characteristics are almost horizontal, indicating a high output resistance for the CB amplifier. Although the collector-to-base voltage is negative in this case (for a PNP), it is common practice to plot the voltage to the right of the origin.

9-9. FORWARD CURRENT TRANSFER RATIO, CB

The graph of collector current against collector-base voltage is a series of practically horizontal lines with I_E the controlling factor. I_B is very small compared with I_C; therefore, $I_C \simeq I_E$. This leads to the definition of forward current transfer ratio, CB, and a similar discussion to that for the CE is possible where a distinction is made between dc and ac current gain.

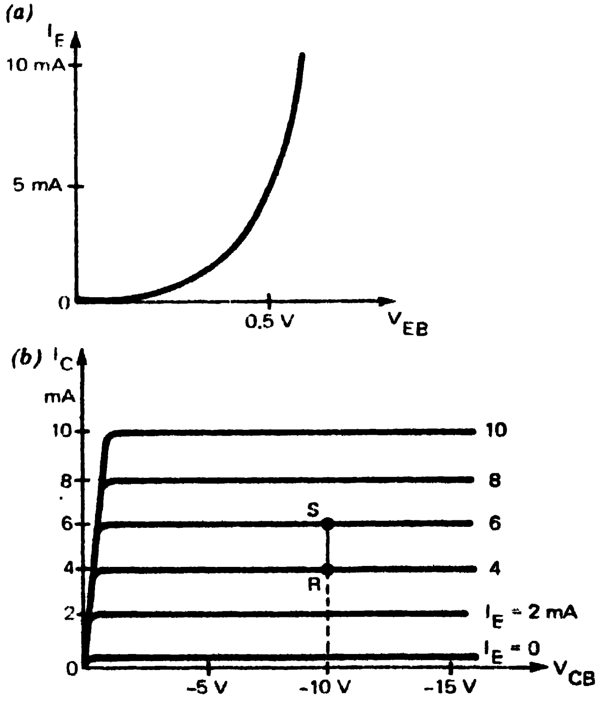


Fig. 9-12. Characteristic curves for a transistor used in the CB connection. (a) Input characteristics. (b) Output collector characteristics.

Dc Current Gain

The static forward current transfer ratio in a CB is the ratio $\frac{I_C}{I_E}$ and may be designated as α, α_{dc} or h_{FB}. That is:

$$\alpha = \alpha_{dc} = h_{FB} = \frac{I_C}{I_E},\ (\alpha = \text{alpha}) \tag{9-7}$$

where I_C is the dc collector current flowing as a result of a dc current I_E in the emitter.

Again, a distinction is sometimes made between the large signal current gain, α, and the dc current gain, α_{dc}. But, since the cutoff current I_{CBO} (Section 9-11) is usually negligible, there is little difference between the two.

Evaluating α at S in Fig. 9-12(b) yields the following:

$$\alpha = \frac{I_C}{I_E} = \frac{5.88\ \text{mA}}{6\ \text{mA}} = \frac{5.88 \times 10^{-3}}{6 \times 10^{-3}} = 0.98.$$

Ac Current Gain

The small-signal gain or forward-current transfer ratio in a CB is the ratio $\frac{\Delta I_C}{\Delta I_E}$ at a given collector-to-base voltage. It is referred to as α_{ac} or h_{fb}.

$$\alpha_{ac} = h_{fb} = \frac{\Delta I_C}{\Delta I_E},\ V_{CB} \text{ constant.} \tag{9-8}$$

Evaluating this on the curves of Fig. 9-12(b) at $V_{CB} = 10$ V would be difficult, but may turn out as follows:

When $I_E = 4$ mA, $I_C = 3.92$ mA.

When $I_E = 6$ mA, $I_C = 5.88$ mA.

Thus $$\alpha_{ac} = h_{fb} = \frac{\Delta I_C}{\Delta I_E} = \frac{5.88 - 3.92\ \text{mA}}{6 - 4\ \text{mA}}$$

$$= \frac{1.96\ \text{mA}}{2\ \text{mA}} = 0.98.$$

Once again, it is observed that the ac and dc current gains are equal and, although α does depend upon I_E (just as β depends upon I_C), we will usually be safe in using the symbol α without specifying ac or dc current gain.

[h_{fb} is also used in the analysis of small-signal amplifiers in Chapter 14, along with h_{fe}. Note the subscripts. h_{fb} is called the h parameter for the small-signal-*forward* current transfer ratio in a common-*base*, while h_{fe} refers to the *forward* current transfer ratio in a common-*emitter*.]

Alpha, α, is still called the current *gain* in a CB even though it is less than one. It should be evident then that a CB amplifier exhibits a current gain, A_i, less than one. However, since the current in the output can flow through a load resistor that is larger than the input resistance, it is possible to have a voltage gain, A_v, and therefore a power gain $A_p = A_vA_i$. In fact, it is the ability of the device to *transfer* current from a low resistance input to a large *resistance* output that originated the name "transistor" as a contraction of the phrase "*trans*fer res*istor*."

9-10. RELATION BETWEEN α AND β

It was noted in Section 9-9 that it was difficult to determine α accurately, since the difference between I_E and I_C was very small. However, there must surely be some relationship between α and β, since they both refer to the same transistor collector current.

Consider the following equations:

$$I_E = I_B + I_C \tag{9-9}$$

$$\alpha = \frac{I_C}{I_E} \tag{9-10}$$

$$\beta = \frac{I_C}{I_B}. \tag{9-11}$$

From (9-9) $I_B = I_E - I_C$. (9-12)

Substituting this in Eq. 9-11,

$$\beta = \frac{I_C}{I_B} = \frac{I_C}{I_E - I_C}. \tag{9-13}$$

Divide both numerator and denominator of the right hand side by I_E.

$$\beta = \frac{\frac{I_C}{I_E}}{\frac{I_E}{I_E} - \frac{I_C}{I_E}}. \text{ But } \frac{I_C}{I_E} = \alpha.$$

$$\therefore \beta = \frac{\alpha}{1 - \alpha}. \tag{9-14}$$

Transposing this yields

$$\alpha = \frac{\beta}{1 + \beta}. \tag{9-15}$$

Thus, if β is determined accurately (which is easily done), α can be calculated.

Say β was measured to be 49. Then $\alpha = \frac{49}{1 + 49} = 0.98$. Therefore, for a transistor to have a large β it must have an α close to one (unity).

9-11. COLLECTOR CUTOFF CURRENT, I_{CBO}

It can be seen from the collector characteristics of the CB that I_C decreases as I_E is reduced until at $I_E = 0$ the collector current is very small. The reverse current that does flow through the base-collector junction (due to the *drift* of minority carriers, which are available at room temperature) when the emitter is open is called the reverse saturation current I_{CBO}, the current from collector to *b*ase with the emitter *o*pen. It is also referred to as the collector cutoff current I_{CO} and is strongly temperature-dependent.

Since $\alpha = \frac{I_C}{I_E}$, then $I_C = \alpha I_E$. When the cutoff current is included

$$I_C = \alpha I_E + I_{CBO}.$$

This equation describes the collector characteristics for the common-base.

If the collector current equations for a CE and CB are compared, their cutoff currents are seen to be significantly different:

$$\text{CB: } I_C = \alpha I_E + I_{CBO} = \alpha I_E + I_{CO} \tag{9-16}$$

$$\text{CE: } I_C = \beta I_B + I_{CEO} = \beta I_B + (\beta + 1) I_{CO}. \tag{9-17}$$

That is, the CE equation has been put in terms of the cutoff current I_{CO} in a CB amplifier for comparison purposes. (This is explained in Section 9-15.) Evidently the leakage current in a CE is β times larger than in a CB. Since I_{CO} is temperature-dependent, then $(\beta + 1)I_{CO}$ means a large change in collector current in the CE circuit when temperature changes.

For example, a 2N708 silicon transistor has an I_{CO} @ 25°C = 0.025 μA and I_{CO}@ 150°C = 15 μA, an increase of 600 times.

If this transistor were used in a CE circuit and β was approximately 50, then the leakage current I_{CEO} would be 1.25 μA @ 25°C and 0.75 mA @ 150°C. What this means in terms of the stability of the transistor will be covered later, but it can be stated now that the CE characteristic curves have a tendency to move upward significantly with increase in temperature and may cause thermal runaway unless guarded against. The CB circuit is much less temperature-dependent.

In silicon transistors, I_{CO} is usually negligible except at very low values of bias current, but a germanium transistor's leakage current is appreciably larger. In general, I_{CO} can be taken to increase 7 to 8% per °C increase and doubles with a temperature change of 10°C.

9-12. TRANSISTOR TESTERS

Many of the parameters mentioned in the preceding sections may be measured on a transistor tester, of which many versions are available. A picture of a portable transistor tester is shown in Fig. 9-13. It is designed for testing power type transistors (h_{FE}), signal type transistors (h_{fe}), and diodes.

The dependency of I_{CEO} on temperature may readily be seen by first checking the transistor's I_{CEO} while the transistor is at room temperature and then observing the meter deflection as the transistor is held tightly between the fingers.

Some transistor testers have the ability to measure both ac and dc β and are often designed for in-circuit operation. That is, the transistor does not have to be removed from the circuit in which it is connected, in order for tests to be made.

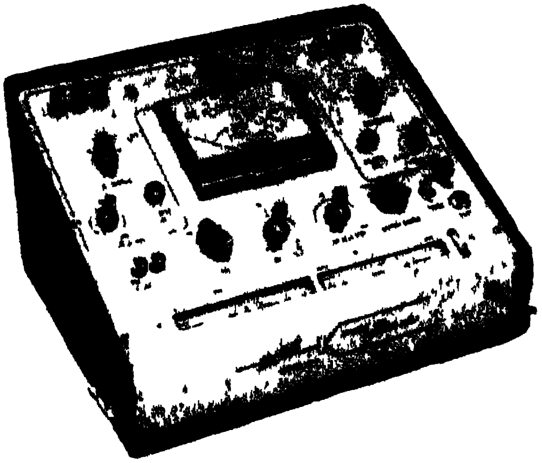


Fig. 9-13. Transistor tester. (Courtesy of Hickok Electrical Instrument Co.)

9-13. PHYSICAL EXPLANATION OF CB AND CE AMPLIFYING ACTION

It was stated that a CE circuit exhibited current gain due to the base current controlling the degree to which the base-emitter junction was turned on, which in turn controlled the collector current. This explanation could be used in this case since the base-emitter junction was in *series* with the collector supply voltage and collector-base junction. See Fig 9-14.

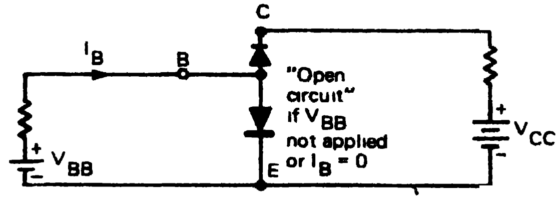


Fig. 9-14 Showing how collector current depends upon the base-emitter junction bias

Unless the base-emitter junction were forward-biased, V_{CC} could not supply any current from collector to emitter (or from emitter to collector if electron flow is followed), except for the leakage current.

Common-Base Action

In a CB circuit however, the simple two-diode analogy is somewhat misleading Refer to Fig 9-15.

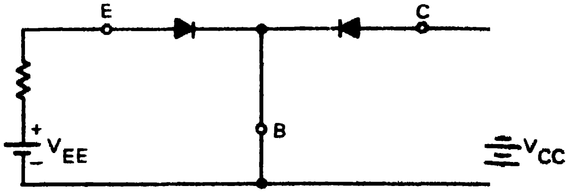


Fig 9-15 Common-base biasing using the two-diode analogy.

It seems that if the collector-base junction is simply a "leaky diode," then V_{CC} could supply just as much current in the base-collector circuit when the emitter is open as when it is forward-biased. However, let us examine the operation under normal forward-biasing conditions with reference to the PNP slices, as shown in Fig. 9-16.

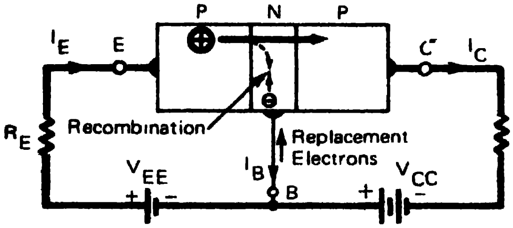


Fig. 9-16. CB biasing showing PNP slices and base current control.

The forward biasing of the emitter-base junction allows the series combination of V_{EE} and V_{CC} to inject holes (the majority carriers in the P-type) through the emitter toward the base. The doping of the base is less than that of the emitter or collector, so that not many recombinations of holes with electrons take place. A factor that aids this situation is the narrowness of the base (width is only 1 to 2% of total length of transistor), so that the holes are not in the base sufficiently long for many recombinations to take place. Roughly 98% of the holes injected to the base from the emitter are swept across the base and into the collector. This collector current is relatively independent of the voltage from collector to base, so the characteristic curves for a CB are essentially flat. The remaining 2% requires a supply of electrons for replenishment purposes supplied by V_{EE}. (This constitutes a small base current I_B, which in conventional current designation is shown flowing *downward* from the base and can be thought of as the 2% of the holes.) Therefore, if $I_E = 10$ mA then $I_C = 9.8$ mA and $I_B = 0.2$ mA. That is, the base current is of the same order of magnitude as in a CE amplifier. Note: If V_{EE} is not present to supply the replacement electrons the emitter-base junction is not forward-biased, but is open, and V_{CC} has no closed path by which it can inject holes from the emitter into the collector. If V_{CC} acts alone, without V_{EE}, the only effect is to reverse-bias the collector-base junction, widening the depletion region and allowing only the minority carriers to drift across the junction, as shown in Fig. 9-17.

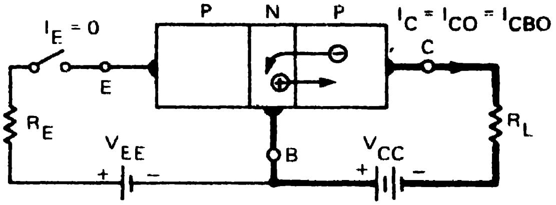


Fig. 9-17. Showing how no forward bias in a CB allows only a small collector leakage current, I_{CBO}.

This is the reverse saturation current or cutoff current I_{CO} and is very small. Actually then, the collector-base diode is not a "leaky diode" as first suggested and used in the CE circuit. The only reason it will allow such a relatively large current to flow across the junction when V_{EE} is present from emitter to base is that the holes acquire sufficient energy in the emitter to be swept across the narrow N-type base to the heavily doped P-type collector. Due to the *weak* doping of the narrow base, the transistor—from emitter to collector—looks essentially like one P-type material. With V_{EE} omitted, however, even the narrow N-type base can develop a sufficient barrier to keep the current down to an almost negligible amount.

Common-Emitter Action

Consider now the PNP transistor used in the CE configuration, as in Fig. 9-18. The forward biasing of the base-emitter junction allows

V_{CC} to inject holes from the emitter through the narrow, weakly doped base to the collector. Roughly 98% of the injected holes are collected by the collector, constituting a large reverse current and giving the *effect* of a "leaky diode." The remaining 2% of holes recombine with electrons in the N-type base. V_{BB} serves to replenish these electrons and constitutes the base current I_B. If the availability of replacement electrons is reduced by reducing V_{BE} and hence I_B (that is, reducing the forward bias), then not as many holes will be ejected from the emitter toward the collector, with the result that collector current I_C decreases. Thus, it is possible for a small base current to control a large collector current.

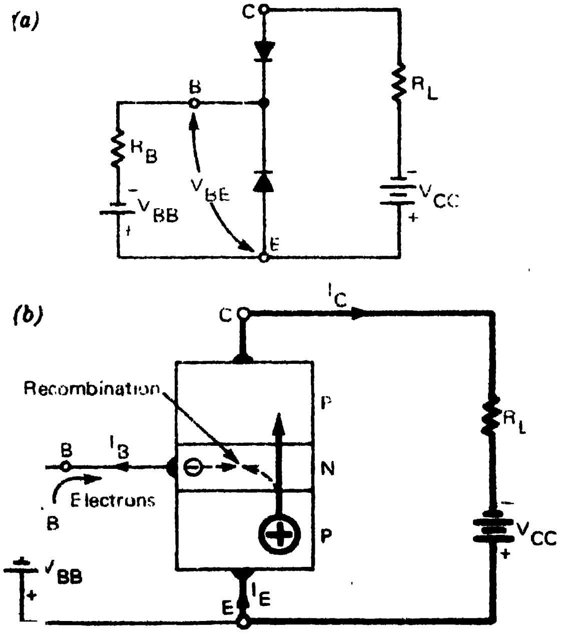


Fig. 9-18. (a) Two-diode model showing biasing in a PNP CE circuit. (b) Showing movement of holes and electrons when forward biased.

Why is the collector current dependent upon V_{CC}? As V_{CC} is made to increase the reverse bias, the space charge width between collector and base tends to increase (this means the barrier extends into the base), with the result that the *effective width of the base decreases.* This dependency of base width on collector-to-emitter voltage is known as the *Early effect.* What this means is that a narrower base will reduce the number of recombinations of holes with electrons, so that the collector current will increase perhaps from 98% to 98.5% of I_E. This is effectively increasing α and β and accounts for the upward slope of the CE characteristic curves. Now consider the CE with the base open, as illustrated in Fig. 9-19.

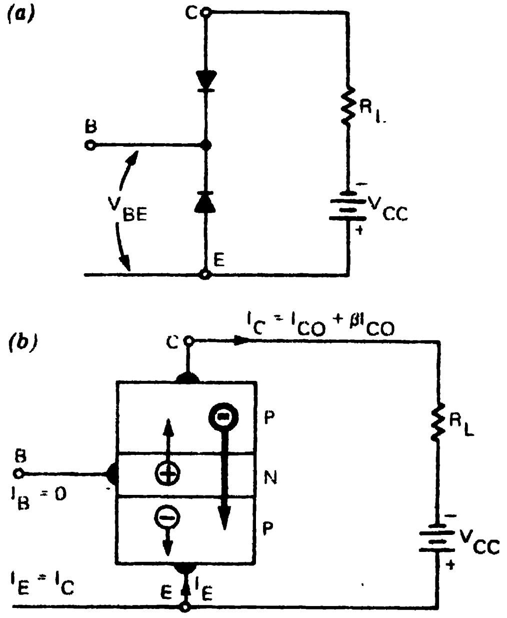


Fig. 9-19. (a) Two-diode model of CE with open base. (b) Showing movement of holes and electrons in a CE with no forward bias.

It is to be noted that the leakage current is not due just to the flow of minority carriers across the collector-base junction as in the CB circuit, I_{CO}, but there is also a movement of electrons across the base-emitter junction.

This electron flow has the effect of slightly *forward biasing* the base-emitter junction (typically 30 mV exists for V_{BE} when the base is open), as if base current were being supplied equal to I_{CO}. But this potential difference will allow the movement of holes from the emitter to the collector in the reverse direction as far as the collector-base junction is concerned. This hole current will give rise to an *additional* collector current of

βI_{CO}, since $\beta = \frac{I_C}{I_B}$ and I_B is simulated by I_{CO}. That is, $I_C = \beta I_B = \beta I_{CO}$.

Hence, the total leakage current flowing through the transistor with the base open is $I_{CO} + \beta I_{CO} = (\beta + 1)I_{CO}$. As has been mentioned before, this is referred to as I_{CEO}.

9-14. REDUCTION OF CE LEAKAGE CURRENT TO I_{CO}

Since the CE circuit is the most popular for switching circuits, and since this leakage current *could* be large enough to be detrimental in some applications, the question may arise how can this leakage current be reduced?

If the minority carriers in the P-type collector (the electrons) are prevented from crossing the base-emitter junction, the forward-biased condition will not exist, and the *only* leakage current will be I_{CO}. To prevent minority current flow, a positive voltage is placed on the N-type base with respect to the emitter, so that electrons will go out through the base and not across the junction.

This is a condition of reverse bias across the base-emitter junction, as shown in Fig. 9-20. In this condition, $I_E = 0$, so the transistor is truly cut off. It requires approximately 0.1 V reverse bias for V_{BE} (whether PNP or NPN) in a germanium transistor (which is often used in high-speed switching circuits) and 0 V in a silicon transistor. That is, it is only necessary to short the base to the emitter in a silicon transistor to cut the leakage current down to I_{CO}. This is because when the collector current is in the order of I_{CO}, α and β are very nearly zero in a silicon transistor, so that

$$I_C = I_{CO} + \beta I_{CO} \simeq I_{CO}. \tag{9-18}$$

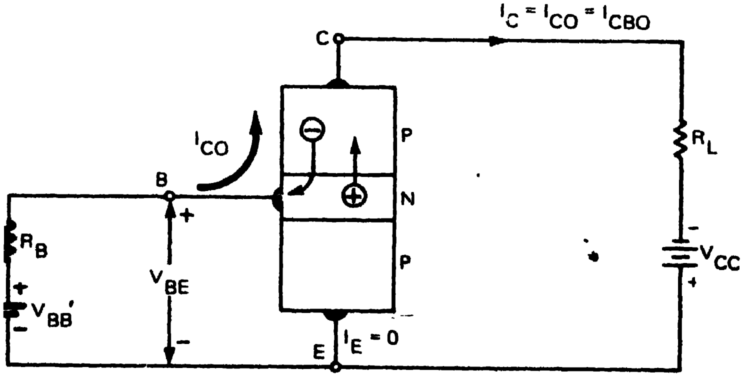


Fig. 9-20. Showing how reverse bias from base-to-emitter reduces the leakage current in a CE to I_{CO}.

In a germanium transistor however, β is still approximately 10 even at low collector currents, as depicted in Fig. 9-21.

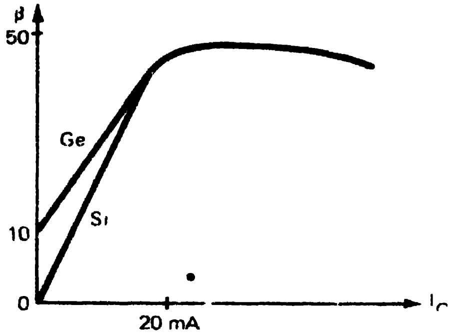


Fig. 9-21. Graph showing how β varies with low values of collector current for Ge and Si transistors.

9-15. COMMON-COLLECTOR AMPLIFIER

One basic transistor circuit configuration remains, called the grounded or common collector. It also goes by the name emitter follower and is the counterpart of the cathode follower in vacuum-tube circuitry. It is essentially the same as the common-emitter circuit with the output taken in the emitter lead instead of the collector, as shown in Fig. 9-22. The same explanation of operation holds true as for the common emitter circuit, as far as electron and hole motion is concerned. As for a cathode follower, the common collector (since the collector is at ground level for ac signals) shows no voltage gain. That is, A_v is less than unity, but a current gain A_i exists and thus a power gain A_p. A more detailed look at the emitter follower will be taken under the subject of negative feedback in Chapter 17.

9-16. PNP AND NPN TRANSISTOR AMPLIFIERS

Most of the preceding circuits involved PNP transistors. If an NPN were substituted then *all* polarities of voltage supplies would have to be reversed. The explanations for operation would still hold with holes replaced by electrons.

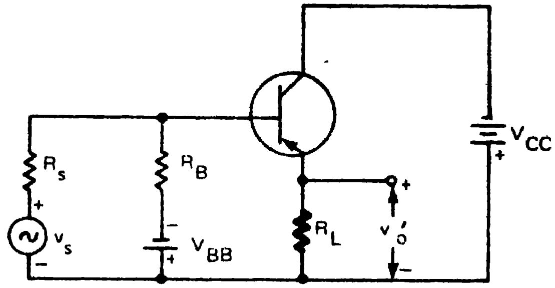


Fig. 9-22. Basic common-collector or emitter-follower amplifier.

9-17. TRANSISTOR CONSTRUCTION

The earliest transistors were of the point-contact type, or "catwhisker" construction, but, because of their instability and the difficulty of analyzing them, they have been replaced by the grown-junction transistors. Many construction methods exist, each giving a particular property to the transistor.

In the *grown-junction* type, a single crystal is drawn from a melt of silicon or germanium while the impurity concentration is changed by adding materials to give N-type or P-type as required. The grown crystal is then sliced into a large number of small-area devices, and contacts are made to each region, as shown in Fig. 9-23.

In *alloy-junction* transistors, two small "dots" of a P-type or N-type impurity element are placed on opposite sides of a thin wafer of N-type or P-type semiconductor material, respectively. The whole structure is raised for a short time to a temperature high enough to melt the impurity into the base material. The alloying of the impurity with the original wafer produces the PN junctions, to which leads are attached as shown in Fig. 9-24 The collector is made larger than the emitter to collect the emitter current as soon as possible and prevent the current from diffusing into the whole base.

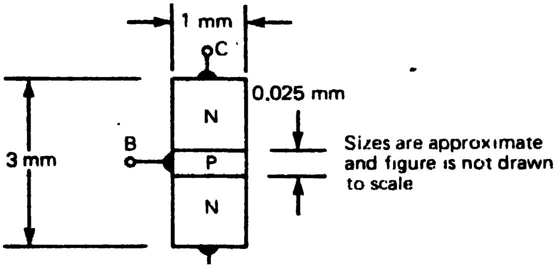

1 mm
C
0.025 mm
N
B
3 mm
P
N
Sizes are approximate and figure is not drawn to scale

Fig. 9-23. Schematic of a grown-junction type of transistor. (From Electronic Devices and Circuits *by Millman and Halkias. Copyright (c) 1967 by McGraw-Hill, Inc. Used with permission of McGraw-Hill Book Company.)*

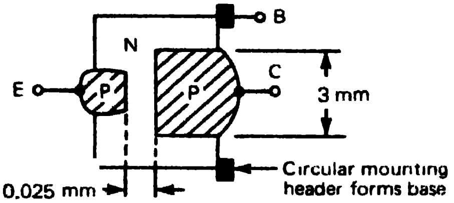

B
N
C
E
P
P
3 mm
0.025 mm
Circular mounting header forms base

Fig. 9-24. Diagram of an alloy-junction type of transistor. (From Electronic Devices and Circuits *by Millman and Halkias. Copyright (c) 1967 by McGraw-Hill, Inc. Used with permission of McGraw-Hill Book Company.)*

If, instead of having a well defined boundary for the base between emitter and collector, the impurities are allowed to diffuse into the base, a *drift-field* transistor results, which speeds current flow and reduces capacitive effects, making the transistor capable of higher-frequency operation.

High-frequency operation required the development of new construction techniques. The *diffusion* technique led to the silicon planar transistor, which is closely related to integrated circuit construction. Consider the construction of an NPN silicon planar transistor, a cross-section of which is shown in Fig. 9-25.

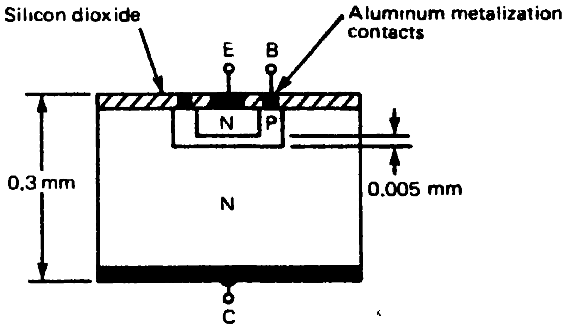

Silicon dioxide
Aluminum metalization contacts
E
B
N
P
0.3 mm
N
0.005 mm
C

Fig. 9-25. Cross-section of an NPN silicon planar transistor. (From Electronic Devices and Circuits *by Millman and Halkias. Copyright (c) 1967 by McGraw-Hill, Inc. Used with permission of McGraw-Hill Book Company.)*

Starting with a block of N-type silicon, the base-collector junction area is photoetched by photolithographic techniques, and the P-type base is applied by a gaseous diffusion masking process. The emitter is then diffused onto the base, and the whole structure is covered over by a final layer of silicon oxide, through which contacts are made with the base and emitter. The passivating action of the oxide layer reduces surface problems and provides low leakage currents.

Another type of transistor is the *mesa*, which is essentially a planar transistor that has been etched to reduce the collector area at the base collector junction, leaving a "mesa" or flat-topped peak, as illustrated in Fig. 9-26.

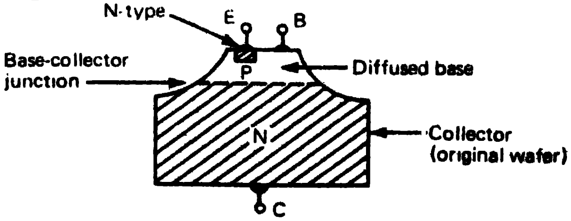

N-type
E
B
Base-collector junction
P
Diffused base
N
Collector (original wafer)
C

Fig. 9-26. Sketch of a high-frequency mesa type of transistor.

The mesa structure is inherently rugged, has large power dissipation capability, and can operate at very high frequencies.

If a separate collector-contact diffusion process is used—called epitaxial growth—to reduce the electrical series resistance in the collector, the result is an *epitaxial-mesa transistor.*

Various combinations of the previous methods exist, such as a drift field in the mesa and planar structures that result in reduced base-lead resistance. Also, relatively light doping in the collector region results in high collector-breakdown voltages and low collector-junction capacitance.

9-18. IDENTIFYING THE TRANSISTOR LEADS

The symbols for the PNP and NPN transistors are shown in Fig. 9-27.

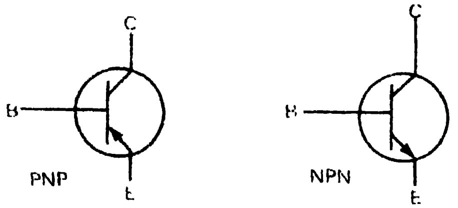


Fig. 9-27. Symbols used for PNP and NPN transistors.

The arrow on the emitter serves to distinguish a PNP (arrow-in) from the NPN (arrow out). However, the physical transistor cannot be identified as being one or the other without the use of a transistor manual to identify what the number on the transistor means, such as a 2N109 being a PNP type.*

The base, collector, and emitter may be found, for four different mountings, by referring to Fig. 9-28. These do not include power transistors, in which one lead is often the flange of the case in which the transistor is mounted.

9-19. SUMMARY

(1) A transistor may be compared loosely with a triode in that the emitter may be thought of as a nonheated cathode, the collector as a plate, and the base as a grid.

(2) Since a transistor consists of either PNP or NPN slices, a very simplified model for biasing purposes comprises two diodes "pointing inward" or "outward" respectively.

(3) The transistor input characteristic is essentially that of a forward-biased junction, with some dependence upon the transistor's output voltage.

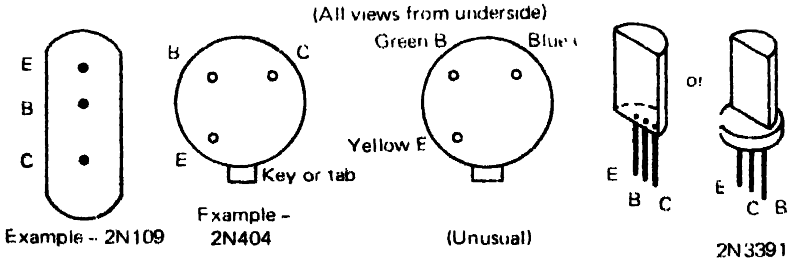


Fig. 9-28. Showing how to identify leads in four different transistor mountings.

---

*It is also possible to identify the transistor type by making resistance checks between base and emitter with an ohmmeter whose battery polarity is known.

(4) The collector or output characteristics for a transistor used in the CE connection have the collector current vs. collector voltage curves sloping upward slightly with base current as the controlling parameter.

(5) The CE ac current gain, $\beta_{ac} = h_{fe} = \frac{\Delta I_C}{\Delta I_B}$ (V_{CE} = constant), while $\beta_{dc} = h_{FE} = \frac{I_C}{I_B}$.

(6) The permissible operating area of a transistor's characteristics are defined by $I_{C,max}$, the maximum dissipation hyperbola, and $V_{(BR)CEO}$.

(7) For a transistor used in the CB connection the collector characteristics consist of practically horizontal lines of I_C vs. V_{CB}, with I_E as the controlling parameter.

(8) The CB ac current gain, $\alpha_{ac} = h_{fb} = \frac{\Delta I_C}{\Delta I_E}$ (V_{CB} = constant), while $\alpha_{dc} = h_{FB} = \frac{I_C}{I_E}$, and is less than unity.

(9) $\alpha = \frac{\beta}{1 + \beta}$ and $\beta = \frac{\alpha}{1 - \alpha}$.

(10) The reverse saturation current in a CB connection = $I_{CBO} = I_{CO}$, while in a CE connection it is $I_{CEO} = (\beta + 1)I_{CO}$.

(11) In either CB or CE action, majority carriers are swept across the narrow, lightly doped base to the collector. This current is controlled by a relatively small base input current that is the result of the forward-biased base-emitter junction, which accounts for the amplifying action in a transistor.

QUESTIONS

Q9-1. Distinguish between the following notations: i_c, I_C and i_C; $I_E = I_C + I_B$ and $i_E = i_C + i_B$; V_{BB} and V_{CC}.

Q9-2. In what manner are the base-emitter and collector-base junctions biased in a PNP transistor when used in an amplifier?

Q9-3. If a transistor used in the CE connection has its collector voltage increased, what will happen to the base current if the base voltage is held constant? Why?

Q9-4. Using the two-diode analogy explain why the base-emitter junction has to be forward biased to provide collector current. If the base-emitter junction is open, what current will flow and how is this designated?

Q9-5. What is β or h_{FE} a measure of? How does it differ from β_{ac} or h_{fe}? When is there any significant difference between the static and dynamic values of beta?

Q9-6. What equation describes approximately the CE collector characteristics?

Q9-7. What three quantities determine the permissible operating area of a transistor's collector characteristics?

Q9-8. Sketch a common-base amplifier circuit with an NPN transistor and indicate clearly the polarities of the supply voltages. Do the same with a common-emitter amplifier circuit with an NPN transistor. Can you see why a CE amplifier may be preferred over a CB as far as supply voltages are concerned?

Q9-9. What do the almost horizontal lines in the CB characteristic curves indicate concerning the output resistance of a CB amplifier?

Q9-10. Compare the magnitudes of current in the input characteristics of a CE and CB in Figs. 9-4 and 9-12, and suggest which amplifier has the higher input resistance.

Q9-11. Distinguish between: $h_{FB}$and h_{fb}; h_{fe} and h_{fb}; α and β.

Q9-12. Describe in your own words the origin of the name "transistor."

Q9-13. Describe two methods for obtaining h_{fb} or α. Which is the more accurate?

Q9-14. Consider a transistor in the CB connection with zero emitter current. What is the designation used to describe the collector current under this condition, and what effect would a reduction in temperature have on this current?

Q9-15. How much larger is the CE leakage current compared with the CB leakage current for a given transistor? Which one, if either, is affected more by temperature?

Q9-16. What problems can you see in using an in-circuit transistor tester?

Q9-17. Why can the collector-base PN junction be thought of as a "leaky diode" in a CE configuration as in Fig. 9-14?

Q9-18. The degree to which the base-emitter junction is forward biased clearly controls the collector current in a CE circuit. Explain in your own words how the forward biasing of the emitter-base junction also controls collector current in a CB circuit.

Q9-19. If a larger number of recombinations of electrons and holes takes place in the base, does this mean more or less base current?

Q9-20. What is the *Early effect*? What does it account for in the transistor's characteristic curves?

Q9-21. Explain in your own words why I_{CEO} $(\beta + 1)I_{CO}$. ($I_{CO} = I_{CBO}$.)

Q9-22. What methods are used to reduce the emitter current to zero in a CE circuit for silicon and germanium transistors? Why is it simpler for a silicon transistor?

Q9-23. How is it possible for power gains to occur in CB and CC amplifiers if the first has $A_i < 1$ and the second has $A_v < 1$?

Q9-24. Explain what the effect would be on the size of base current if a transistor having the construction shown in Fig. 9-24 were operated with the emitter and collector interchanged.

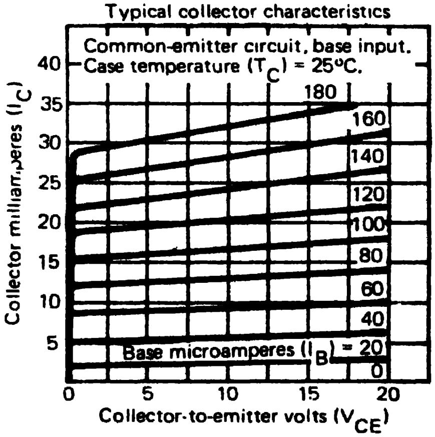
Q9-25. How does a mesa-type transistor differ from a planar transistor, and what advantages does it have?

PROBLEMS

P9-1. For the given CE characteristic curves of a silicon NPN transistor, determine $\beta(h_{FE})$ at the following points:

(a) $I_B = 20\ \mu A$, $V_{CE} = 15$ V.
(b) $I_B = 100\ \mu A$, $V_{CE} = 15$ V.
(c) $I_B = 180\ \mu A$, $V_{CE} = 15$ V.

Sketch a graph of h_{FE} vs. I_C.



Curves for P9-1 through 4

(Courtesy of RCA Corp.)

P9-2. Using the given characteristic curves, determine β_{ac} (h_{fe}) at $V_{CE} = 15$ V and for various collector currents from 2 mA to 32 mA. Sketch a curve of h_{fe} vs. I_C and compare with P9-1.

P9-3. Repeat P9-1 using $V_{CE} = 2.5$ V instead of 15 V. Sketch the curve of h_{FE} vs. I_C on the same axes as in P9-1 and comment on the results.

P9-4. Repeat P9-2 using $V_{CE} = 2.5$ V instead of 15 V. Sketch the curve of h_{fe} vs. I_C on the same graph as in P9-2. Which is more constant over the transistor's characteristics, h_{fe} or h_{FE}?

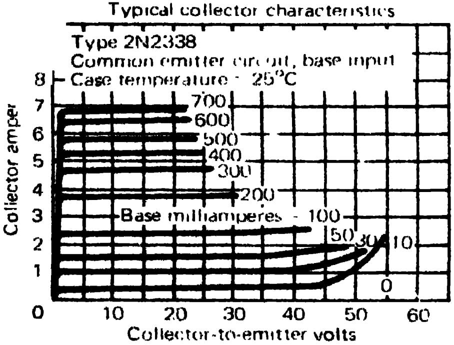
P9-5. (a) A germanium transistor's collector current is 51 mA when the base current is 0.4 mA. If $h_{FE} = 125$, what is the transistor's collector cutoff current I_{CEO}?

(b) Why would the method you used be impractical for determining a silicon transistor's I_{CEO}?

P9-6. Using the average value for h_{FE} from P9-1 and Eq. 9-15, obtain the average value for h_{FB}.

P9-7. Using the average value for h_{fe} from P9-2 and Eq. 9-15, obtain the average value for h_{fb}.

P9-8. Given the CE characteristic curves for a 2N2338 silicon power transistor, indicate the permissible area of operation for the following maximum values: $I_C = 7.5$ A, $V_{CEO} = 40$ V, Power dissipation = 150 W at case temperatures up to 25°C.



Curves for P9-8 through 11

(Courtesy of RCA Corp.)

P9-9. Using the 2N2338 characteristic curves, determine h_{FE} at $V_{CE} = 10$ V and base currents of 10 mA, 100 mA and 600 mA. Sketch a graph of h_{FE} vs. I_C.

P9-10. Using the 2N2338 curves, determine h_{fe} at $V_{CE} = 10$ V and in the region of base currents of 0 to 10 mA, 50 to 100 mA and 600 to 700 mA. Sketch a graph of h_{fe} vs. I_C. How does this compare with the small-signal transistor in P9-2?

P9-11. Find the average value of h_{FE} from P9-9 and determine the value of h_{FB}. Repeat for h_{fb} from P9-10.

P9-12. A transistor has its emitter current increased from 15 to 20 mA. This caused an ammeter in the base lead to increase its reading from 0.32 mA to 0.48 mA.

(a) Calculate h_{fb}.

(b) Determine h_{fe} in two ways.

P9-13. A 40396 is a germanium transistor used in complementary symmetry amplifiers. It has a collector cutoff current $I_{CBO} = 14$ μA at room temperature and $h_{FE} = 50$.

(a) What is the collector current when the base current is 0.2 mA?

(b) Assuming h_{FE} does not increase with temperature, what would be the new collector current if the transistor's temperature *rises* 50°C? (I_{CBO} doubles with a temperature change of 10°C.)

(c) Assuming the transistor's h_{FE} varies with temperature according to the curve in Fig. 9-7, recalculate the collector current for a 50°C rise in temperature.

10

THE COMMON-BASE AMPLIFIER

10-0. INTRODUCTION

This chapter discusses the graphical analysis of current and voltage gains in a common-base amplifier. Graphical analysis uses the input and output characteristic curves of the transistor and is applicable to large-signal and power amplifiers, where the device is operated over large voltage and current variations. This type of analysis provides a powerful method of understanding the operation of an amplifier even if the reader does not use it as a design tool.

Since the bipolar transistor is a current-operated device (as contrasted with a tube, that is a voltage-controlled device), the effect of the internal resistance of the source is considered in obtaining overall amplifier gains. The input and output waveforms are also examined, as are the power relationships and operating efficiencies.

10-1. ANALYSIS OF THE COMMON-BASE AMPLIFIER

Given the common-base circuit input and output characteristics as shown in Fig. 10-1, it is required to analyze the operation of the circuit and determine graphically the voltage, current, and power gains.

Note that R_L is the load resistor that develops the output voltage from the transistor. C_2 is used to couple the ac component of the output to R_o, which represents the input impedance of the next stage of amplification or whatever load is connected to the amplifier.

The Dc Load Line

The first step is to draw the dc load line on the output characteristic curves. This is done in a similar manner to that used in Chapter 2 for the diode analysis. In this case, the equation that relates the variables in the collector circuit is

$$v_{CB} \quad i_C R_L - V_{CC} \qquad (10\text{-}1)$$

which represents a number of coordinates of i_C and v_{CB} for a given V_{CC} and R_L.

Let us consider two extreme coordinate points.

(1) One end of the load line is where $i_C = 0$. $v_{CB} = -V_{CC}$. In this case, $v_{CB} = -V_{CC} = -16$ V. Plot this point, $i_C = 0$, $|v_{CB}| = 16$ V. This often called the cutoff point.

(2) The other end of the dc load line is located where $v_{CB} = 0$.

$$\therefore 0 = i_C R_L - V_{CC}$$

$$i_C R_L \quad V_{CC}$$

$$\frac{V_{CC}}{R_L}$$

In this case, $i_C = \frac{16\text{ V}}{400} = \frac{16\text{ V}}{0.4\text{ K}} = \frac{160\text{ V}}{4\text{ K}}$ $= 40$ mA. Plot this point, $i_C = 40$ mA, v_{CB} 0 V.

A line drawn connecting these two points is called the dc load line, since it represents all the possible combinations of i_C and v_{CB} for the given load resistor R_L. The point at which the line intersects the characteristic curves near the collector current axis is referred to as the saturation point. This is the condition of maximum current through the transistor and minimum-voltage drop across it for this value of load.

The Operating Point

The dc load line evidently crosses a number of curves, not just one as in the case of the diode analysis in Section 2-1. The dc operating point,

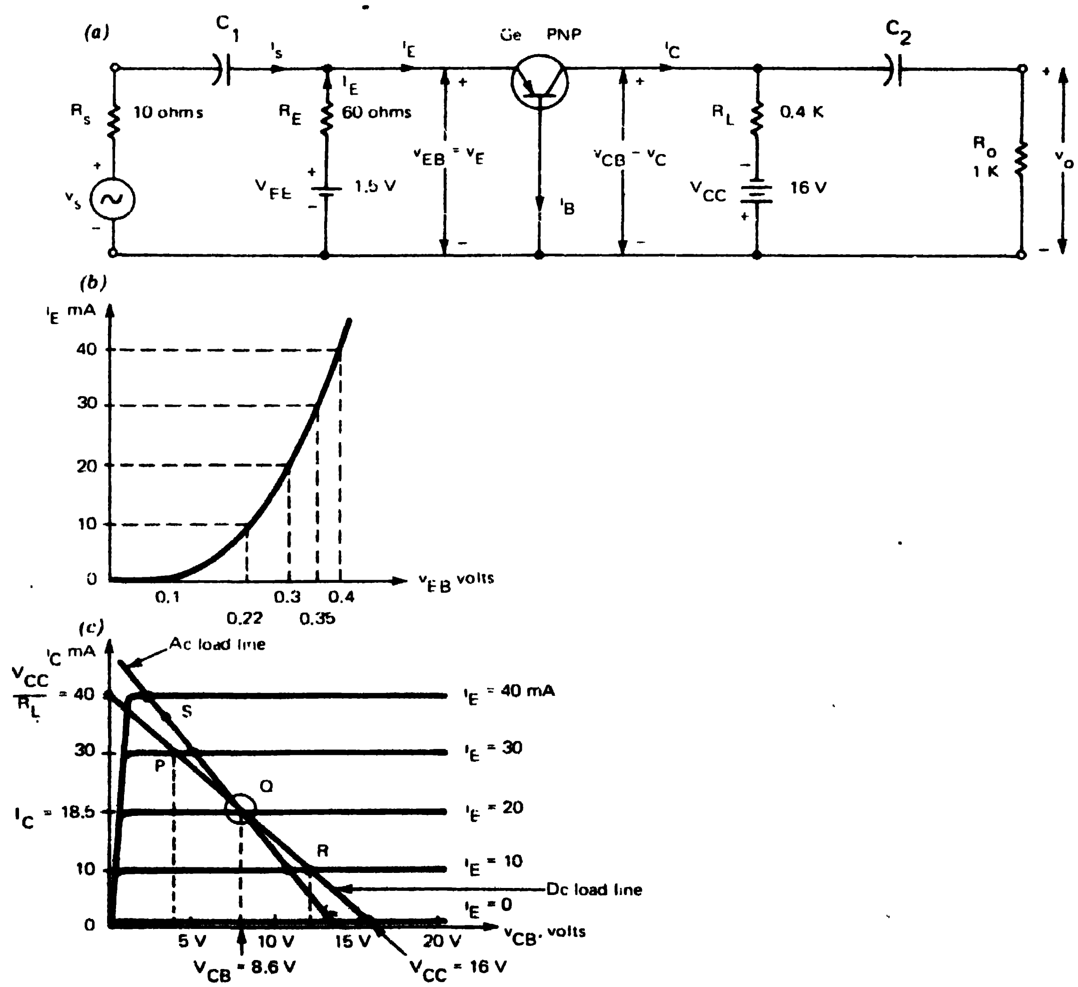

(a)
C1
is
iE
Ge PNP
iC
C2
Rs
10 ohms
RE
60 ohms
vEB = vE
vCB = vC
RL
0.4 K
Ro
1 K
vo
vs
VEE
1.5 V
iB
VCC
16 V
(b)
iE mA
40
30
20
10
0
0.1
0.22
0.3
0.35
0.4
vEB volts
(c)
iC mA
VCC/RL = 40
30
iC = 18.5
10
0
Ac load line
S
P
Q
R
iE = 40 mA
iE = 30
iE = 20
iE = 10
iE = 0
Dc load line
5 V
10 V
15 V
20 V
vCB, volts
VCB = 8.6 V
VCC = 16 V

Fig. 10-1. (a) Common base circuit to be analyzed.
(b) Emitter-base input characteristics.
(c) Collector characteristics, showing the dc and ac load lines.

also referred to as the quiescent point, is determined by finding the dc value of emitter current I_E that will be supplied by $V_{EE} =$ 1.5 V. Write an equation for the input circuit:

$$V_{EE} = I_E R_E + V_{EB}. \qquad (10\text{-}2)$$

Now in a germanium transistor the approximate forward bias voltage from emitter to base is in the order of 0.3 V (approximately 0.6 V for silicon).

$$V_{EE} = I_E R_E + 0.3$$

$$1.5 = I_E \times 60 + 0.3$$

$$60 I_E = 1.5 - 0.3$$

$$\therefore\ I_E = \frac{1.2}{60} = 0.02 \text{ A} = 20 \text{ mA}.$$

The dc operating point can now be located at the intersection of $I_E = 20$ mA and the dc load line (Q point).

$I_C = 18.5$ mA, $V_{CB} = 8.6$ V.

Note: Normally the value used for V_{EE} is larger (say 12 V), in which case V_{EB} is negligible, and so $V_{EE} \approx I_E R_E$ and $I_E \approx \dfrac{V_{EE}}{R_E}$.

Referring to Fig. 10-1(c), it is evident that the purpose of V_{EE} and R_E is to bias (place) the no-signal operating point of the amplifier in the middle of the transistor's characteristic curves or in an area where the curves are relatively equally spaced. Then, when a signal is applied by varying the emitter current, the collector current and voltage will vary instantaneously as shown by *moving along the load line* to provide an output relatively free from distortion. Note that C_1 has been added to prevent the dc bias V_{EE} from supplying current to the input source, v_s, which would just be a waste of power. However, the ac source is still coupled to the emitter through C_1 as long as the reactance of C_1 is chosen to be low compared with the input resistance to the transistor at the lowest frequency under consideration.

Dc Input Resistance

Some idea of the dc input resistance to the transistor may be obtained from the input characteristic curve shown in Fig. 10-1(b), which shows how emitter current varies with emitter-base voltage. This curve is not generally directly available from the transistor manuals, but what is often available is the *collector* current against base-emitter voltage graph. Since the emitter current is only about 2 to 5% larger than the collector current (different only by the base current) this curve can be used for our purposes.

Locate $I_E = 20$ mA and draw a horizontal line. The intersection with the curve gives $V_{EB} = 0.3$ V. (This voltage compares exactly with the assumed value used in determining the emitter bias current. This will not always be the case, but a value of 0.4 V for example would make such a small difference in actual emitter bias current — especially for larger values of V_{EE} — that no correction would be made.)

$$\text{Dc input resistance to transistor} = \frac{V_{EB}}{I_E} \qquad (10\text{-}3)$$

$$= \frac{0.3 \text{ V}}{20 \text{ mA}} = 15 \text{ ohms}.$$

Application of Input Signal—Current Gain

To show the effect of loading an amplifier, assume first that the load R_o is disconnected. Also assume that v_s is applied in such a way that the source current i_s makes the emitter current vary 20 mA peak-to-peak from the dc quiescent value of 20 mA. Thus, i_E will vary between a high of 30 mA and a low of 10 mA. The resulting variation in the output circuit for a one-cycle variation at the input can be obtained by moving *along the dc load line* from Q to P representing a quarter of a cycle, back to Q for half a cycle and completing the cycle by moving down to R and then back to Q. Reading off the corresponding values of i_C and v_C may yield:

$I_{C,max} = 28$ mA $V_{C,min} = 4.8$ V for

$I_{E,max} = 30$ mA

$I_{C,min} = 9$ mA $V_{C,max} = 12.4$ V for

$I_{E,min} = 10$ mA

Thus $I_{c,peak\text{-}to\text{-}peak}$ through $R_L = 19$ mA. This resulted from a change at the input of $I_{e,peak\text{-}to\text{-}peak} = 20$ mA.

$$\therefore \text{ Current gain, } A_i = \frac{I_{c,p\text{-}p}}{I_{e,p\text{-}p}} = \frac{19 \text{ mA}}{20 \text{ mA}} = 0.95.$$

As is expected, the ac current gain in a CB amplifier is less than one.

Voltage Gain

To obtain the voltage gain, refer to the input characteristic:

For an extreme of $i_E = 30$ mA, v_E must be 0.35 V.

For a low of $i_E = 10$ mA, v_E must be 0.22 V.

$$\therefore V_{e,peak\text{-}to\text{-}peak} = 0.35 - 0.22 = 0.13 \text{ V}$$

$$\therefore \text{ Voltage gain } A_v = \frac{V_{o,p\text{-}p}}{V_{e,p\text{-}p}} = \frac{V_{c,p\text{-}p}}{V_{e,p\text{-}p}}$$

$$= \frac{12.4 - 4.8}{0.13}$$

$$= \frac{7.6}{0.13} = 58.5.$$

Power Gain

Hence, the power gain, $A_p = A_v A_i = 58.5 \times 0.95 = 55.6$. Power gain in dB = 10 log A_p = 10 log 55.6 = 10 × 1.75 = 17.5 dB, which is fairly typical of the gain one can expect in a CB amplifier except for larger load resistors.

The Ac Load Line

Now consider connecting a load, $R_o = 1$ K ohm, to the amplifier. Assuming the reactance of C_2 is negligible, R_o and R_L will be in parallel with each other as far as the ac components of voltage and current are concerned, constituting an equivalent load, R_e.

$$R_e = \frac{R_L R_o}{R_L + R_o} = \frac{0.4 \text{ K} \times 1 \text{ K}}{0.4 \text{ K} + 1 \text{ K}} = 0.286 \text{ K ohms}$$

When a variation in emitter current now occurs, the swing in collector current and voltage must be along a load line representing an ac load resistance of 286 ohms instead of 400 ohms. This demands that a new line be drawn through the same operating point as before, but what should be its slope? Consider the equation in the output circuit that represents the load line:

$$v_{CB} = i_C R_L - V_{CC}. \tag{10-4}$$

Solving for i_C,

$$i_C = \frac{1}{R_L} \times v_{CB} + \frac{V_{CC}}{R_L}. \tag{10-5}$$

Comparing this equation with the standard equation for a straight line of slope m and y intercept c,

$$y = mx + c, \tag{10-6}$$

it can be seen that the load line has a slope given by $\frac{1}{R_L}$, and an intercept on the i_C axis of $\frac{V_{CC}}{R_L}$.

You will recall that a line that slopes from the upper left to the lower right (as does our dc load line) has a negative slope, whereas the above seems to imply a positive slope. The discrepancy arises from the fact that we are dealing with a PNP transistor in which the voltage from collector to base is normally negative and yet in drawing our graph of collector characteristics we referred to the voltage as positive. In fact, transistor manuals show collector voltage and current as negative in PNP transistors and positive in NPN transistors. We have not followed this convention, but apart from this instance we will not be unduly concerned, since the correct phase relationships and polarities are also obtained with our method as shown later.

The question is, how do we draw a line having a slope given by $\frac{1}{R_L}$, or, as in this case, $\frac{1}{286 \text{ ohms}}$? Using $\frac{1}{R_L} = \frac{I}{V}$, the units of slope must be $\frac{\text{amps}}{\text{volts}}$ or amps per volt. Thus slope = $\frac{1}{286}\frac{\text{amps}}{\text{volts}} = 0.0035$ amps per volt = 3.5 mA per volt. This is equivalent to 17.5 mA per 5 volts, which is more accurate to use in plotting the information.

To draw the ac load line, we start at the Q point on the characteristic curves (8.6 V, 18.5 mA) and move to the left 5 volts and vertically upwards 17.5 mA to locate another point S. The coordinates of this point would be:

$$v_{CB} = 8.6 - 5 = 3.6 \text{ V and } i_C = 18.5 + 17.5$$
$$= 36 \text{ mA}.$$

Using this point and the Q point, we draw the ac load line as illustrated in Fig. 10-1(c).

Evidently the ac load line is steeper than the dc load line (because of a smaller load resistance). And it should be clear from the curves that for the same input current conditions a variation in emitter current of 20 mA p-p will provide a smaller swing in output voltage than with no load connected. Thus the voltage gain is reduced, as shown below:

$$A_v = \frac{V_{o,p\text{-}p}}{V_{e,p\text{-}p}} = \frac{V_{c,p\text{-}p}}{V_{e,p\text{-}p}} = \frac{11.2-6}{0.13} = \frac{5.2}{0.13} = 40.$$

However, the current gain is relatively unaffected at 0.95, and the power gain is reduced to 38 or 15.8 dB.

The Ac Input Resistance

The ac input resistance to the transistor itself may be found from:

$$R_{in} = \frac{V_{e,p\text{-}p}}{I_{e,p\text{-}p}} \tag{10-7}$$

$$= \frac{0.13 \text{ V}}{20 \text{ mA}} = 6.5 \text{ ohms}.$$

It will be shown later, with the equivalent circuit, that the ac input resistance to the transistor is actually dependent upon the load resistance R_L as well. However, for a range of R_L up to 10 K ohms, the input resistance is typically 25 ohms, rising to 500 ohms only if R_L is as high as 10 M ohms. The low input impedance to the CB amplifier is one of the characteristics and drawbacks of the CB connection.

10-2. THE AMPLIFIER AS A WHOLE

It should be emphasized that the above calculations for gain are all with respect to the *input* and *output* terminals of the transistor *itself*, and not at the input and output of the amplifier as a whole. The difference is shown in Fig. 10-2. The gains calculated from B to B′ can be very different from those between A and A′, and also can the input and output resistances be different, according to whether one is talking about the input to the transistor itself or the amplifier as a whole.

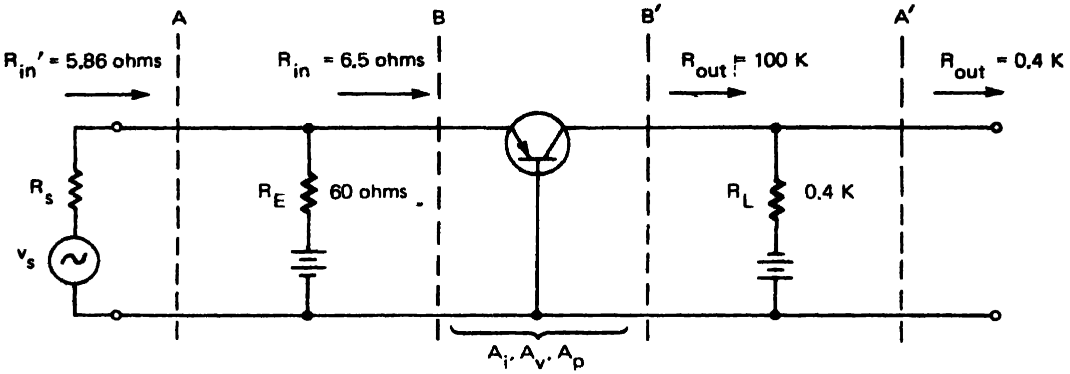


Fig. 10-2. The common-base circuit, showing the difference between the transistor itself and the amplifier as a whole.

Amplifier Input Resistance

It has been shown that the ac input resistance at B in Fig. 10-2 is 6.5 ohms, but when an ac source is connected at A there is an additional 60 ohms resistance due to the emitter bias resistor in parallel with the transistor input. Hence, the overall ac input resistance at A is

$$R_{in} = \frac{60 \times 6.5}{66.5} = 5.86 \text{ ohms.}$$

Amplifier Output Resistance

It may be shown that the output resistance of a common-base connection depends upon the internal resistance R_s of the source and is typically at least 100 K ohms. (This is one of the characteristics of the CB circuit.) But this 100 K ohms is at B′. At A′ the load resistor R_L = 400 ohms is in parallel with 100 K ohms, so that the effective output resistance at the *amplifier* terminals is 400 ohms in parallel with 100 K ohms ≈ 400 ohms. This last statement raises an important point. If the output resistance of the transistor is 100 K ohms would not a larger voltage gain and a higher power transfer occur if R_L were made near 100 K ohms instead of 400 ohms?

This is correct. For R_L = 100 K ohms, the voltage gain would be in the order of 2000, and the current gain would be approximately 0.9, for a power gain of 1800! However, this vast improvement in power gain would only materialize if the load connected to the amplifier is in the order of 1 megohm. If this is not the case, either power gain would be sacrificed or a matching transformer would have to be used. Generally speaking a common-base amplifier has few applications except to match a very low impedance source to a very high impedance load, and in some high-frequency applications.

Amplifier Current Gain

Consider now the comparison of amplifier current gain as opposed to transistor current gain. Refer to Fig. 10-3. Considering only ac quantities, the source will deliver current i_s which is the sum of i_e and i_{ee} (the current that flows through the bias resistor R_E).

$$i_{ee} = \frac{V_{e,p\text{-}p}}{R_E} = \frac{0.13 \text{ V,p-p}}{60 \text{ ohms}} = 2.2 \text{ mA,p-p.}$$

The total source current will be: 20 + 2.2 =

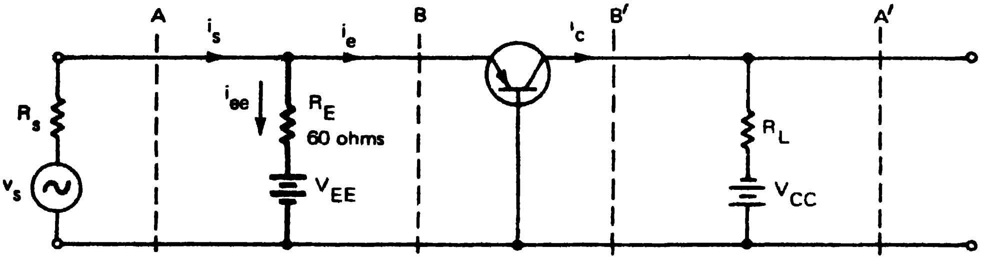


Fig. 10-3. The common-base circuit, showing the total ac current.

22.2 mA,p-p. Amplifier current gain, $A_i = \frac{I_c}{I_s} = \frac{19\text{ mA,p-p}}{22.2\text{ mA,p-p}} = 0.86$. This is to be compared with a transistor circuit current gain of 0.95. The increased loading due to R_E can be reduced by increasing R_E and V_{EE}. For example, if I_E = 20 mA was required with V_{EE} = 12 V, then

$R_E = \frac{V_{EE}}{I_E} = \frac{12\text{ V}}{20\text{ mA}} =$ 600 ohms. What this means is that when the source current is applied a smaller amount of it will flow through the biasing circuit if R_E = 600 ohms compared with the case where R_E = 60 ohms. Thus, the larger resistor *reduces loading* on the input source.

Amplifier Voltage Gain

It was determined that a peak-to-peak voltage of 0.13 V was required from emitter to base to develop the output of 7.6 V peak-to-peak. However, what would be the necessary value of v_s (the open circuit value of the source voltage) if R_s = 10 ohms? Refer to Fig. 10-4, which shows a partial circuit of the input to the transistor.

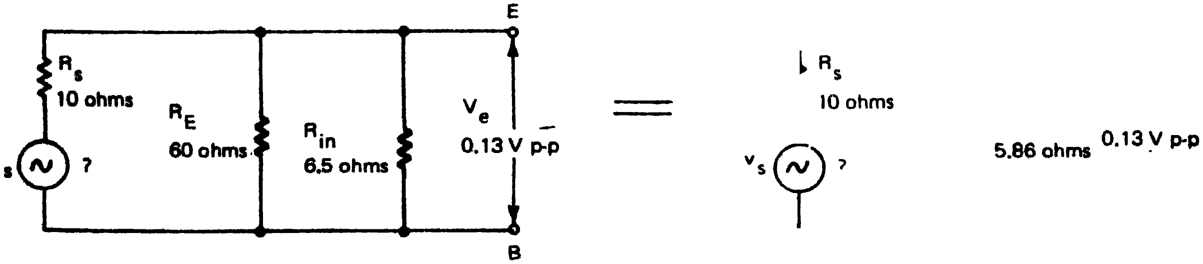


Fig. 10-4. Partial equivalent circuit for ac input signal to the transistor.

By the voltage divider method:

$$0.13 = v_s \times \frac{5.86}{10 + 5.86}$$

$$\therefore 0.13 = v_s \times \frac{5.86}{15.86}$$

$$\therefore v_s = \frac{0.13 \times 15.86}{5.86} = 0.35 \text{ V,p-p.}$$

Actual overall amplifier voltage gain =

$\frac{V_o}{V_s} = \frac{7.6 \text{ V,p-p}}{0.35 \text{ V,p-p}} = 22.$ This compares with a gain of 58.5 from input to output of the transistor.

This same result could have been obtained by knowing that the total source current must be 22.2 mA,p-p, as found in Section 10-2, and that the total ac input resistance, as far as v_s is concerned, is $R_s + R_{in}' = 10 + 5.86 = 15.86$ ohms.

Thus $v_s = i_s R_T = 22.2 \times 10^{-3} \times 15.86 = 0.35$ V,p-p.

Note: If R_s had been 1 ohm instead of 10 ohms, then the open circuit value for v_s would only have to be 0.15 V,p-p, and the overall amplifier gain would be 50. Thus, there is a strong dependence on the internal resistance of the source in determining overall voltage gain.

10-3. POWER RELATIONSHIPS

It was mentioned that graphical analysis is especially appropriate for large-signal amplifiers. In order to obtain a typical value for the efficiency of such an amplifier, let us consider that the amplifier in Fig. 10-1(a) is driven just to cutoff on one extreme and near to saturation at the other, along the ac load line drawn in Fig. 10-1(c). This will require an emitter current swing of 40 mA, p-p, and 0.3 V,p-p, from emitter to base. The corresponding collector variations may be 38 mA,p-p, and an output voltage of 11 V,p-p.

Assuming that the input waveforms are sinusoidal (it will be shown later that the voltage waveform usually is somewhat distorted), the ac input power *to the transistor* is given by:

$$P_i = \frac{V_{e,p\text{-}p}}{2\sqrt{2}} \times \frac{I_{e,p\text{-}p}}{2\sqrt{2}} \tag{10-8}$$

$$\therefore P_i = \frac{0.3 \times 40 \times 10^{-3}}{8} \text{ watts} = 1.5 \text{ mW.}$$

The ac output power delivered to R_L and R_c is given by:

$$P_o = \frac{V_{c,p\text{-}p} \times I_{c,p\text{-}p}}{8} \tag{10-9}$$

$$= \frac{11 \times 38 \times 10^{-3}}{8} \text{ W} = 52 \text{ mW.}$$

$$\text{Power gain, } A_p = \frac{P_o}{P_i} = \frac{52 \text{ mW}}{1.5 \text{ mW}} = 34.7.$$

(It is left to the reader to show that this same value for A_p would be obtained from $A_p = A_v A_i$ for this condition of current drive.) Note that this power gain is not from source to output, since the ac power lost in the bias resistor has not been taken into account.

10-4. EFFICIENCY

As mentioned previously, an amplifier is basically an active device that converts dc power from the battery to ac signal power in the load. The degree to which the device is successful in doing this is measured by the conversion efficiency, which is sometimes called the collector circuit efficiency in a transistor amplifier. This is defined as follows:

$$\eta = \frac{\text{ac signal power delivered to the load}}{\text{dc input power to the collector circuit}}. \tag{10-10}$$

The signal power has already been calculated as P_o above, and is 52 mW. The collector dc input power is the product of the supply voltage V_{CC} and the *average* current that it supplies, I_C.

$$\text{Thus } \eta = \frac{P_o}{V_{CC} I_C} \times 100\%$$

$$= \frac{52 \text{ mW}}{16 \times 18.5 \text{ mW}} \times 100\% = 17.5\%.$$

It will be shown later that the maximum theoretical value for η is 25% for loads that are not transformer-coupled. This value can be closely reached with the proper selection of load line, operating point, and current drive in a transistor amplifier, especially one employing a common-base connection.

10-5. PHASE RELATIONSHIPS IN A CB AMPLIFIER

Consider the input and output current and voltage relations in Fig. 10-5, with an applied signal of 20 mA, p-p into the emitter.
In the output circuit:

$$v_C = v_o = i_C R_L - V_{CC}.$$

With no v_s, $I_C = 18.5$ mA with $R_L = 0.4$ K, V_{CC} 16 V.

$$\therefore v_o = 18.5 \times 10^{-3} \times 0.4 \times 10^3 - 16 = 7.4 - 16 = -8.6 \text{ volts.}$$

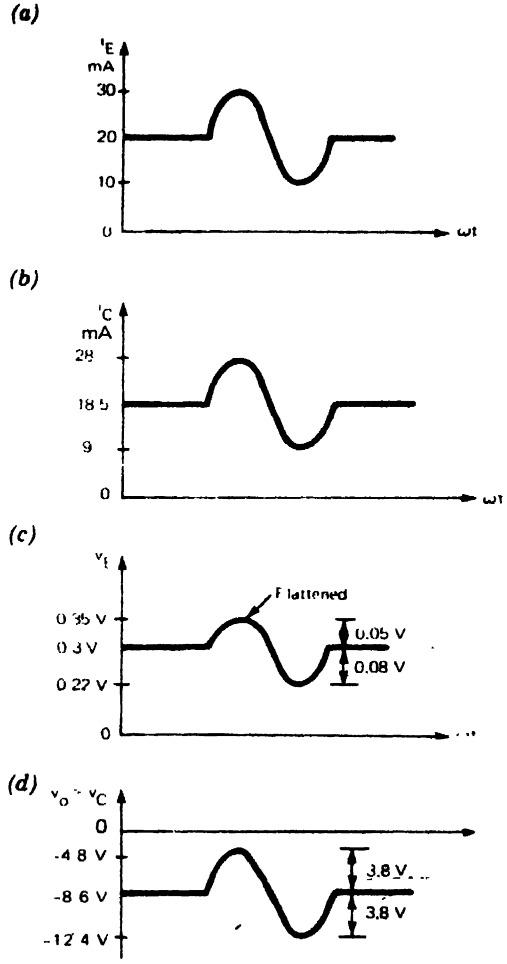


Fig. 10-6. Waveforms of current and voltage in a CB amplifier.

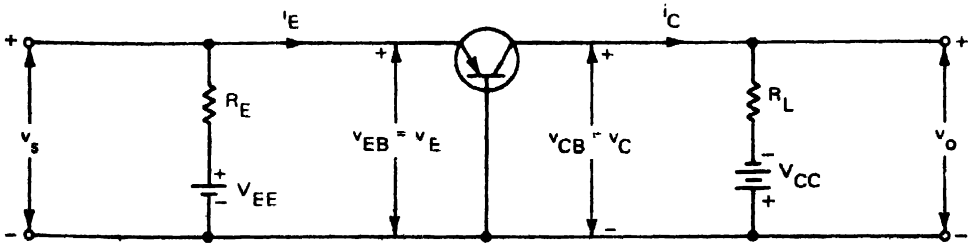


Fig. 10-5. Circuit used to determine phase relationships in a common-base amplifier.

When v_s was applied so that v_{EB} increased from 0.3 V to 0.35 V, i_E increased from 20 mA to 30 mA and i_C from 18.5 mA to 28 mA.

$$\begin{aligned}\therefore v_o &= i_C R_L - V_{CC} = 28 \times 10^{-3} \times 0.4 \times 10^3 \\ &\quad - 16 = 11.2 - 16 = -4.8 \text{ volts.}\end{aligned}$$

When v_E decreased to 0.22 V from 0.3 V, i_E decreased to 10 mA, and i_C to 9 mA.

$$\begin{aligned}\therefore v_o &= i_C R_L - V_{CC} \\ &= 9 \times 10^{-3} \times 0.4 \times 10^3 - 16 \\ &= 3.6 - 16 = -12.4 \text{ volts.}\end{aligned}$$

Thus, the output voltage is *in phase* with the input voltage, as shown in Fig. 10-6.

10-6. INPUT WAVEFORM CONSIDERATIONS

It is to be noted from Fig. 10-6(c) that the waveform of voltage from emitter to base is badly distorted (a 60% greater decrease in voltage from the bias level than the increase above the bias level) when the input voltage v_s is applied. However, the emitter current *is* a sine wave, with the result that the output current and voltage are sinusoidal. The reason for the distortion in v_E lies in the nonlinear relationship between v_E and i_E. On the peak of the positive half of the cycle, the input resistance instantaneously is $\frac{v_E}{i_E} = \frac{0.35 \text{ V}}{30 \text{ mA}} =$ 12 ohms, whereas on the negative half cycle it is $\frac{0.22 \text{ V}}{10 \text{ mA}} = 22$ ohms.

Hence, the voltage from emitter to base will tend to be rounded or flattened on the positive peak due to an additional voltage drop across R_s. This will be the case if the source resistance R_s of v_s is *larger* than the input resistance to the transistor. In fact, v_E may sometimes look like a full-wave rectified sine wave, so acute is the distortion due to the change in input resistance to the transistor on the positive and negative half cycles.

If R_s is *smaller* than the input to the transistor (which is very unusual in a CB circuit), then it is possible for v_E to be sinusoidal but i_E to be nonsinusoidal, with a corresponding increase in distortion at the output. This is quite the opposite of the CE amplifier, whose collector characteristics have a more unequal spacing than do the CB curves, so that a distortion of the current at the input helps to compensate for the distortion that would occu in the output.

10-7. HARMONIC DISTORTION

It can be seen from Fig. 10-6 that the collector current and voltage waveforms have equal positive and negative swings. This implies no even harmonic distortion or at least a negligible amount. Even when the transistor is driven near to its extremes on the output characteristic curves, little distortion results in the common-base amplifier. The distortion can, however, be calculated either by the three- or five-point method, as outlined in Section 7-12 Since distortion is not usually a problem in a properly designed common-base amplifier, we will delay a numerical evaluation of distortion until the common-emitter amplifier is considered.

10-8. SUMMARY

(1) The graphical analysis of a large-signal common-base amplifier involves the following steps:

(a) Draw the dc load line: one end at $v_{CB} = 0$, $i_C = \frac{V_{CC}}{R_L}$; the other at $v_{CB} = V_{CC}$, $i_C = 0$.

(b) Locate the operating point, Q, where $i_E = \frac{V_{EE}}{R_E}$.

(c) Draw the ac load line through the Q point with a slope given by $\frac{1}{R_e}$, where $R_e = R_L \| R_o$.

(d) Obtain the current and voltage gains by reading off the corresponding values of current and voltage from the ac load line for a given emitter current drive.

(2) The current gain is less than one, but the voltage gain may be anywhere from 50 to 2000.

(3) When the amplifier as a whole is considered, the gains and resistances calculated for the transistor itself must be modified by taking into account bias, load, and source resistances.

(4) The common-base amplifier is used primarily to match a very low source impedance (approximately 20 ohms) to a high load impedance (100 K ohms and up).

(5) The ac output power may be calculated from

$$P_o = \frac{V_{c,p\text{-}p} \times I_{c,p\text{-}p}}{8}.$$

(6) The conversion or collector circuit efficiency is a measure of how successfully the active device converts dc power to ac signal power and may be close to the maximum value of 25% in a CB amplifier.

(7) The output voltage is in phase with the input voltage.

(8) Distortion of the input voltage waveform may occur due to the very low input resistance of the amplifier.

(9) Very little harmonic distortion is generated in the output of a CB amplifier except at very high values of load resistance.

QUESTIONS

Q10-1. Draw a common-base circuit that utilizes an NPN transistor. Indicate the polarities of the emitter and collector supply voltages. The following questions apply to this circuit:

(a) If the transistor is silicon, how does the input characteristic curve differ from that shown in Fig. 10-1(b)?
(b) If the transistor is silicon, is there any difference in the collector characteristics shown in Fig. 10-1(c)?
(c) What is the equation relating collector voltage to collector current? Is there any change in drawing the dc load line for this circuit compared with that in Fig. 10-1(c)?
(d) What is the equation relating dc emitter current and voltage?
(e) How would the dc input resistance with a silicon transistor compare with that for a germanium transistor?

Q10-2. (a) What does a dc load line represent?
(b) May this line be used when an ac signal is applied?
(c) When is it necessary to draw an ac load line?

Q10-3. (a) Explain why a loaded voltage gain is smaller than the open circuit voltage gain of a CB amplifier.
(b) What does a heavier load mean as regards the slope of the ac load line?

Q10-4. If an ac load line is to be drawn with a slope of 2 mA per volt, explain how this might be done.

Q10-5. (a) What causes the input resistance to the overall amplifier to be lower than the input resistance to the transistor itself?
(b) How may this effect be reduced?

Q10-6. What is the output resistance of the overall CB amplifier usually determined by?

Q10-7. How does a low value of emitter bias supply voltage affect the current gain in a CB amplifier?

Q10-8. What should be the internal resistance of the source to make the overall amplifier's voltage gain close to the transistor voltage gain? What drawback would this type of operation have?

Q10-9. (a) What would be the waveform of the emitter-base voltage if a sine-wave signal were applied that had zero source resistance?

(b) If this signal were relatively large, what would be the effect on the waveform of emitter current? Why? What detrimental effect would this have?

Q10-10. (a) How is it possible for a distorted sinusoidal voltage from emitter to base to provide an undistorted sine wave output voltage?

(b) What condition is necessary to provide this?

Q10-11. Why does Eq. 10-8 not always accurately determine the input signal power? Under what conditions will it be accurate?

Q10-12. If the output and input voltages from a CB amplifier are used to provide a Lissajous figure on an oscilloscope, what figure should result and what phase relationship does this indicate?

Q10-13. A CB amplifier is being used as a voltage amplifier. Its collector load resistance is 10 K ohms providing an input resistance of 25 ohms to a signal that has a source resistance of 50 ohms. In an attempt to increase voltage gain, a technician increases R_L to 1 M ohm but finds that an increase in distortion occurs in the output. What is the reason for the increase in distortion? (Hint: Consider the effect of the increase in input resistance as in Section 10-1.)

PROBLEMS

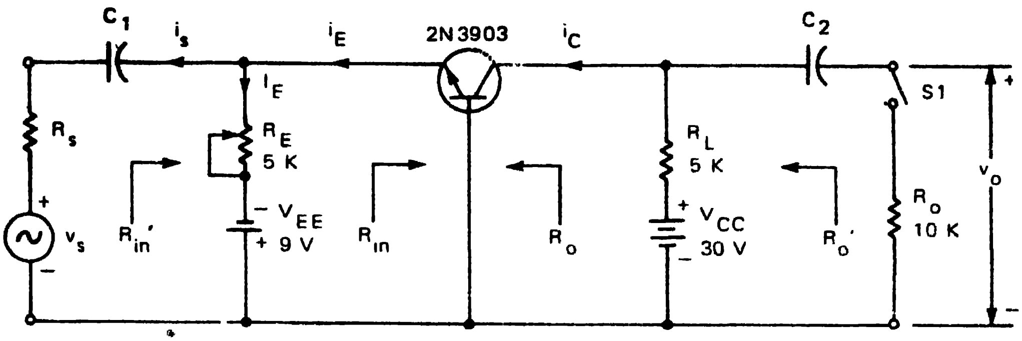
P10-1. Given the CB amplifier circuit shown on the next page, with the 2N3903's characteristic curves, determine the following:

(a) The dc load line.
(b) The value to which R_E should be adjusted to provide an emitter bias current of 3 mA.
(c) The value of V_{CB} and I_C at the operating point in (b).
(d) The emitter-base bias voltage for the operating point in (b).
(e) The dc input resistance to the transistor.
(f) The open circuit output voltage (S-1 open) when a 4 mA peak-to-peak emitter current is applied.
(g) The peak-to-peak emitter-base voltage to provide the input current in (f).
(h) The open circuit voltage gain from (f) and (g).
(i) The peak-to-peak output current for the input current in (f).
(j) The current gain from (i).
(k) The ac input resistance to the transistor from (g).
(l) The power gain from (h) and (j).
(m) The ac load line when S1 is closed.
(n) The output voltage with a 4 mA, p-p input current (S1 closed).
(o) The loaded voltage gain.
(p) The current gain.
(q) The power gain in decibels.

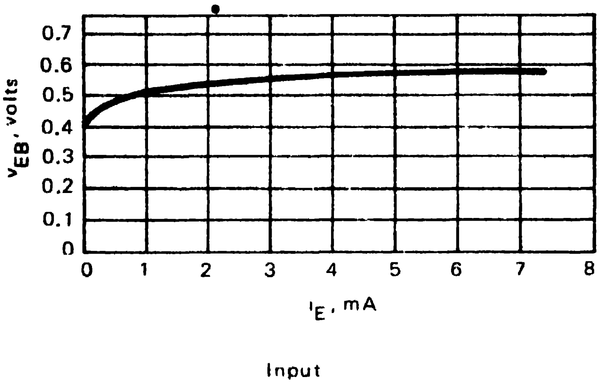
P10-2. Repeat parts (f) through (q) assuming a 6 mA peak-to-peak emitter current is applied (same conditions and operating point as in parts (a) through (e)).

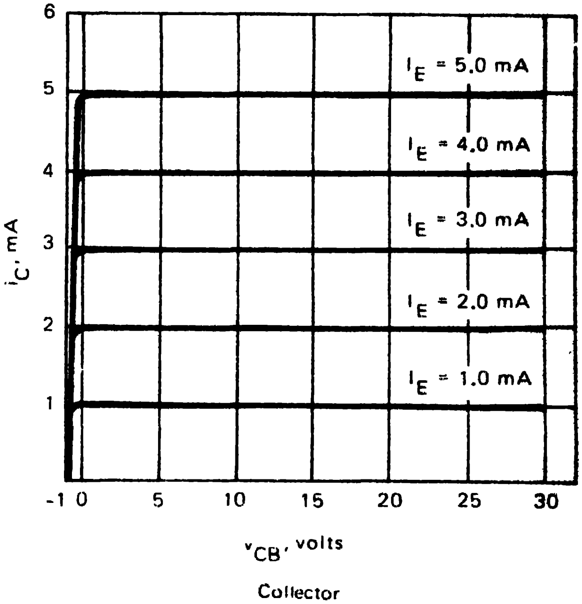
P10-3. Refer to the amplifier and conditions of P10-1. Determine:

(a) The overall ac input resistance to the amplifier, R_{in}'.
(b) The overall amplifier output resistance, assuming the transistor's output resistance



Circuit for problems P10-1 through P10-9.





Characteristic curves of the 2N3903 silicon NPN transistor.

is 100 K ohms.

(c) The overall amplifier current gain with S1 open. Is there any bias resistor loading problem here?

(d) The overall amplifier voltage gain with S1 open and:

(1) R_s : 0 ohms.
(2) R_s : 10 ohms.
(3) R_s : 50 ohms.

P10-4. Repeat P10-3, but use the conditions of P10-2, that is, a 6 mA peak-to-peak signal current.

P10-5. Using the given CB circuit with S1 open and the conditions in P10-1 in parts (a) through (i), calculate:

(a) The ac input power to the transistor, using Eq. 10-8.
(b) The ac output power, to R_L, using Eq. 10-9.
(c) The power gain as a ratio and in decibels. Compare with P10-1(d).
(d) The dc input power to the collector circuit from the V_{CC} supply.
(e) The conversion efficiency using Eq. 10-10.

P10-6. Repeat P10-5 using 6 mA peak-to-peak (instead of 4 mA), as in P10-2. Why is the conversion efficiency higher than in P10-5(e)?

P10-7. Sketch current and voltage waveforms similar to those in Fig. 10-6, showing current and voltage values for the conditions in P10-1, parts (a) through (i).

P10-8. Repeat P10-7, but use the conditions in P10-2, parts (a) through (i), with a 6 mA p-p input current.

P10-9. Assume the CB circuit shown has an ac input resistance of 20 ohms when it is biased at I_E = 1 mA. What will be the shape of the emitter voltage and current waveforms with the following signals applied: (Sketch, showing the dc and peak-to-peak values. Use the input characteristic curve.)

(a) v_s = 0.1 V, p-p, with R_s = 1 ohm.
(b) v_s = 5 V, p-p, with R_s = 1 K ohm.
(c) v_s = 20 V, p-p, with R_s = 10 K ohms.

Which of the above three will cause the least distortion in the output?

(Note: The above waveforms could be obtained accurately using the techniques of Section 2-2 if the input characteristic extended sufficiently high in voltage. However, an approximation of the waveform will be sufficient to show the dependence of distortion on R_s.)

11

THE COMMON-EMITTER AMPLIFIER

11-0. INTRODUCTION

The common-emitter amplifier is the most popular of all three amplifier connections, since it enjoys both a current and a voltage gain, with the result that it exhibits the greatest power gain. Since the large-signal graphical analysis of the common-emitter circuit is very similar in method to that for the common-base, this will not be covered in the same detail. But the analysis does aid in understanding the operation of the amplifier.

In this chapter we will also determine the maximum theoretical conversion efficiency of a class A direct-coupled resistive load amplifier and the harmonic distortion resulting from different values of source resistance.

11-1. GRAPHICAL ANALYSIS OF A COMMON-EMITTER CLASS A AMPLIFIER

Example 11-1. **Given the circuit in Fig. 11-1 and the input and output characteristics in Fig. 11-2, determine the current, voltage, and power gains for a 0.4 mA p-p source current, i_s.**

Solution. Notice that this problem involves the same load and source as for the CB amplifier, so that a comparison can be made between the two.

(a) The first step is to draw the dc load line:

$$v_C = V_{CC} - i_C R_L \, .$$

(1) When $i_C = 0$, $v_C = V_{CC} = 16$ V. Plot.

(2) When $v_C = 0$, $i_C R_L = V_{CC}$.

$$\therefore i_C = \frac{V_{CC}}{R_L} = \frac{16}{0.4\text{ K}} = 40\text{ mA. Plot.}$$

(b) Determine the Q point:

At the input $V_{BB} = I_B R_B + V_{BE}$.

Since this is a silicon transistor, V_{BE} will be approximately 0.6 V.

$$\therefore 16 = I_B \times 40 \times 10^3 + 0.6$$

$$I_B = \frac{16 - 0.6}{40 \times 10^3} = \frac{15.4}{40\text{ K}} = 0.385\text{ mA} \approx 0.4\text{ mA}.$$

The intersection of $I_B = 0.4$ mA with the dc load line gives the Q point: $I_C = 20$ mA, $V_C = 8$ V. Note: In this case V_{BE} is much smaller than V_{BB}, so that:

$$I_B \approx \frac{V_{BB}}{R_B} = \frac{16}{40\text{ K}} = 0.4\text{ mA}.$$

Note also, that instead of using a separate supply, V_{BB}, to provide the bias current I_B, R_B could have been connected to V_{CC}, to give a *one battery operation*, as shown in Fig. 11-3(a) and redrawn in Fig. 11-3(b).

This is called fixed bias, and R_B can always be chosen so as to give the required base bias current from the supply voltage V_{CC}. This is one of the greatest advantages of the CE amplifier.

(c) The magnitude of the dc input resistance in a CE can be obtained from the input characteristic.

For $i_B = I_B = 0.4$ mA, $V_B = 0.64$ V. (This is somewhat higher than was assumed, but will make very little difference in the Q point determination.)

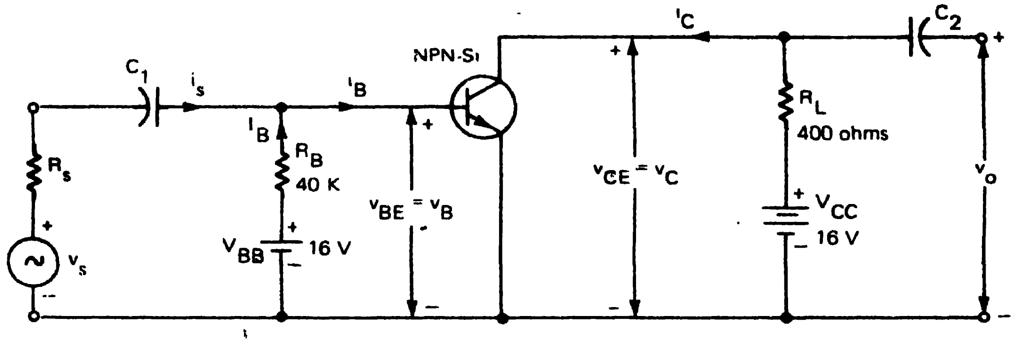

C1
is
iB
IB
RB
40 K
VBB
16 V
NPN-Si
vBE = vB
iC
vCE = vC
RL
400 ohms
VCC
16 V
C2
vo
Rs
vs

Fig. 11-1. Basic common-emitter class A amplifier.

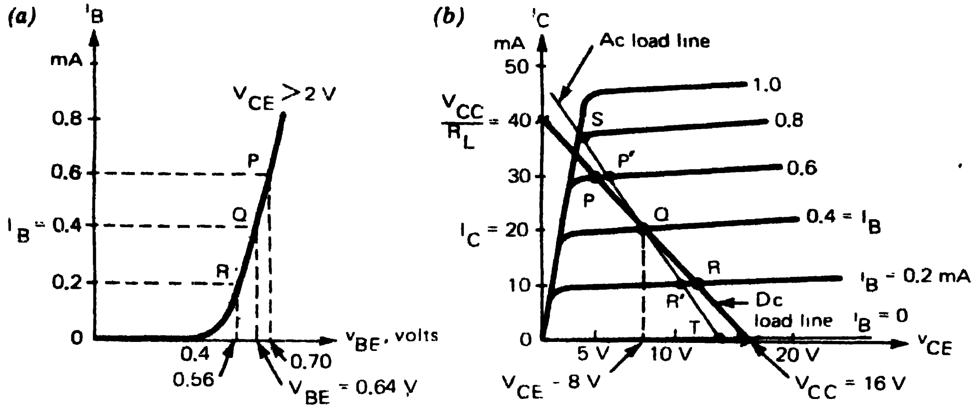


Fig. 11-2. Input and output characteristics for silicon transistor.

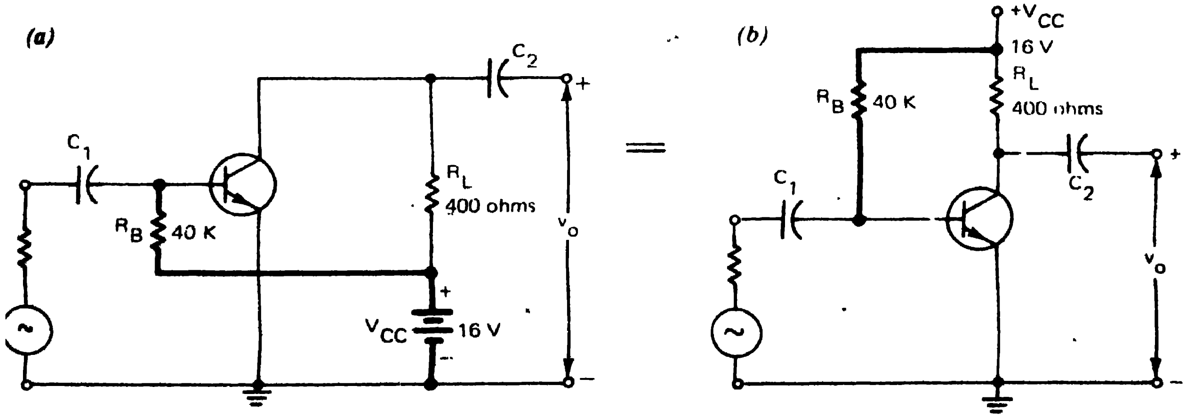

(a)
C1
RB
40 K
RL
400 ohms
C2
vo
VCC
16 V
=
(b)
+VCC
16 V
RB
40 K
RL
400 ohms
C1
C2
vo

Fig. 11-3. Fixed bias, one battery operation of a CE amplifier.

$$R_{in} = \frac{V_B}{I_B} = \frac{0.64\ \text{V}}{0.4\ \text{mA}}\ ;\ \frac{6400}{-} = 1600\ \text{ohms},$$

which is very typical for a CE amplifier, and is certainly higher than the CB.

(d) Now assume that the application of v_s will drive the base current from 0.4 mA to a peak of 0.6 mA and down to a low of 0.2 mA. In other words, the ac component of i_B is 0.4 mA peak-to-peak, or 400 μA p-p.

Reading off the corresponding collector voltages and currents from the load line gives:

P: $I_{C,max} = 30$ mA $\quad V_{C,min} = 4$ V

$I_{B,max} = 0.6$ mA

Q: $I_C = 20$ mA $\quad V_C = 8$ V

$I_B = 0.4$ mA

R: $I_{C,min} = 10$ mA $\quad V_{C,max} = 12$ V

$I_{B,min} = 0.2$ mA

Thus, the ac component of the collector current through R_L is 20 mA peak-to-peak, caused by an input current in the base of 0.4 mA peak-to-peak. Therefore, current gain,

$$A_i = \frac{I_c}{I_b} = \frac{20\ \text{mA, p-p}}{0.4\ \text{mA, p-p}} = \frac{200}{4} = 50.$$

(e) To obtain the voltage gain, refer to v_{BE} on the input characteristic:

P: $I_{B,max} = 0.6$ mA $\quad V_{B,max} = 0.70$ V

Q: $I_B = 0.4$ mA $\quad V_B = 0.64$ V

R: $I_{B,min} = 0.2$ mA $\quad V_{B,min} = 0.56$ V

$\therefore V_{b,\text{peak-to-peak}} = 0.70 - 0.56 = 0.14$ V.

$$\therefore \text{Voltage gain, } A_v = \frac{_{o,p\text{-}p}}{V_{b,p\text{-}p}} = \frac{8\ \text{V}}{0.14\ \text{V}}$$

$$\frac{800}{14} = 57.1\ .$$

Power gain, $A_p = A_v A_i = 57.1 \times 50 = 2855.$

$$\text{Power gain in dB} = 10 \log \frac{P_o}{P_i} = 10 \log A_p$$

$$= 10 \log 2855 = 10 \times 3.455$$

$$= 34.55\ \text{dB},$$

which is 17 dB higher than the CB amplifier.

11-2. INPUT AND OUTPUT RESISTANCE

$$\text{The ac input resistance} = \frac{V_b}{I_b} = \frac{0.14\ \text{V,p-p}}{0.4\ \text{mA,p-p}}$$

$$= \frac{1400}{4} = 350\ \text{ohms}.$$

This is somewhat smaller than the typical value of 1 K ohm, although this is slightly dependent upon the load resistor R_L, dropping to an input of approximately 600 ohms if R_L is as much as 10 M ohms. Hence, there is more isolation in a CE circuit than in a CB. That is, a change in load resistance does not materially change the input resistance in a CE, but there is a larger dependence in the CB circuit.

The output resistance at the transistor terminals can be shown to be typically 40 to 70 K ohms, depending upon the source resistance.

However, it should again be emphasized that gains and input and output resistances will be different for the whole amplifier as compared with the transistor terminals, so that a discussion similar to that for the CB is possible. But, because the current drive into the transistor is much smaller than in the CB, the value of the source resistance R_s does not change the above calculated values very much, unless there is a very high internal resistance in the source compared with the input resistance, R_i. Also, the value of the bias resistor R_B is usually much greater than R_i, so that little additional loading is caused by R_B.

11-3. EFFECT OF ADDING AN AC LOAD

If a load R_o is connected on the output, an ac load line would have to be drawn through the Q point with a slope $= \frac{-1}{R_e}$, where $R_e = R_o \| R_L$, as explained before. The capacitor C_2 serves to block the dc component of voltage from the output, v_o.

For example, if $R_o = 1$ K ohm, then $R_e = 400 \| 1000 \quad 286$ ohms. The slope of the ac load line $\frac{1}{0.286K} = \frac{1 \text{ mA}}{0.286 \text{ V}}$, or 17.5 mA per 5 volts. Drawing this ac load line on Fig. 11-2(b) and letting i_B vary between 0.2 and 0.6 mA gives points P′ and R′. The output voltage is evidently smaller than before, equal to 5.9 V p-p instead of 8 V, so that $A_v = 42.1$ instead of 57.1.

The current gain is, however, substantially unaffected, since the characteristic curves are almost horizontal. In fact there is a slight increase in A_i to 52 so that $A_p = 2190$ (33.4 dB), compared with an unloaded power gain of 2855 (34.55 dB).

11-4. CONVERSION EFFICIENCY OF A CLASS A AMPLIFIER WITH A DIRECT-COUPLED RESISTIVE LOAD

Let us determine the maximum theoretical conversion efficiency of a class A amplifier in which the resistive load is in series with the collector of the transistor. See Fig. 11-4. Let I_c be the rms value of the ac component of the collector current through R_L, and V_{ce} be the rms value of the ac component of the voltage across R_L.

Then the ac power delivered to the load, P_L, is:

$$V_{ce} I_c \text{ watts,} \qquad (11\text{-}1)$$

and the dc power delivered to the collector of the transistor from the power supply (battery), P_B, is:

$$V_{CC} I_{C,Q} \text{ watts.} \qquad (11\text{-}2)$$

Note: $I_{C,Q}$ is the collector current at no-signal input. If this increases and decreases alternately by an amount $\sqrt{2} I_c$ due to a distortion-free sine wave, the average current flowing in the collector is still $I_{C,Q}$, and the reading of the collector dc ammeter will not change as a signal is applied to the base.

If the operating point Q is chosen midway on the load line, this dc input power P_B is dissipated half in the load R_L and half in the transistor when no signal is applied. As an input signal is applied, the power dissipated in the load R_L increases, and the power dissipated in the device decreases. The total output power from the battery is, however, constant; only the division of the power between R_L and the transistor changes.

The conversion efficiency η is given by:

$$\eta = \frac{\text{Total ac power delivered to load}}{\text{Collector dc power supplied by battery}}. \qquad (11\text{-}3)$$

$$\therefore \eta = \frac{P_L}{P_B} \times 100\% = \frac{V_{ce} I_c}{V_{CC} I_{C,Q}} \times 100\%. \qquad (11\text{-}4)$$

In the graphical diagram of the load line, Fig. 11-4(b), let us assume that the characteristic curves originate on the i_C axis ($V_{CE,min}$ is zero). Then:

$$2\sqrt{2} V_{ce} = V_{CC}$$

and

$$I_{C,Q} = \sqrt{2} I_c$$

since $I_{C,min}$ is approximately zero.

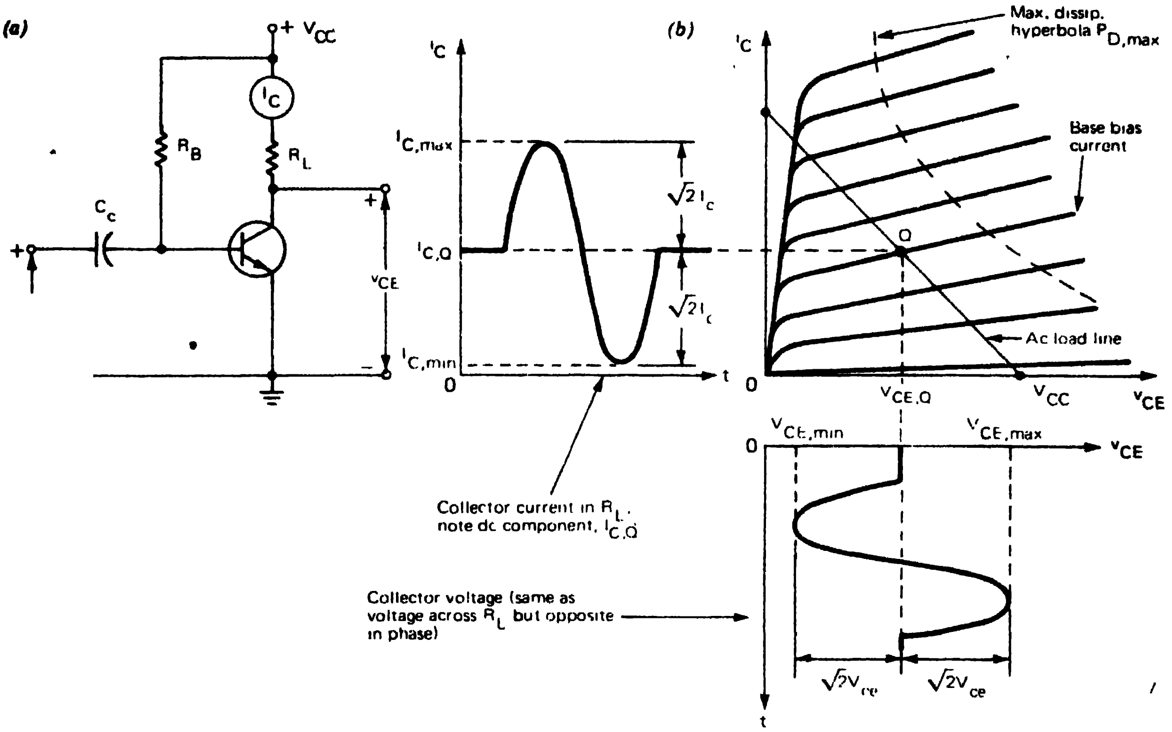

(a)
+ V_{CC}
I_C
R_B
R_L
C_c
+
V_{CE}
−
(b)
I_C
$I_{C,max}$
$I_{C,Q}$
$I_{C,min}$
0
t
$\sqrt{2}I_c$
$\sqrt{2}I_c$
Collector current in R_L, note dc component, $I_{C,Q}$
Collector voltage (same as voltage across R_L but opposite in phase)
Max. dissip. hyperbola $P_{D,max}$
Base bias current
Q
Ac load line
$V_{CE,Q}$
V_{CC}
V_{CE}
$V_{CE,min}$
$V_{CE,max}$
$\sqrt{2}V_{ce}$
$\sqrt{2}V_{ce}$

Fig. 11-4. (a) CE amplifier with direct-coupled resistive load. (b) Output characteristics showing output current and voltage waveforms.

$$\text{Then } \eta = \frac{V_{ce}I_c}{V_{CC}I_{C,Q}} = \frac{V_{ce}I_c \times 100\%}{2\sqrt{2}V_{ce}\sqrt{2}I_c} \qquad (11\text{-}5)$$

$$= \frac{1}{4} \times 100\% = 25\%.$$

This is the optimum efficiency possible under conditions of no distortion and is very nearly reached in a transistor, since $V_{CE,min}$ may be only a fraction of a volt.

The fact that η = 25% means that only a quarter of the dc power supplied by the battery is converted into useful ac power in the load, R_L.

Example 11-2. Determine the conversion efficiency of the circuit in Fig. 11-1 when a 0.8 mA p-p signal current drives the amplifier along the ac load line shown in Fig. 11-2, with a 1 K ohm load.

Solution. Operation between points S and T will result in an 11 V p-p output with a current swing of 36.5 mA p-p. The total ac power delivered to R_L and R_o (assuming negligible distortion) is:

$$P_o = \frac{V_{c,p\text{-}p} \times I_{c,p\text{-}p}}{8} = \frac{11 \times 36.5 \times 10^{-3}}{8} = 50\,\text{mW}.$$

The dc input power supplied by the battery to the collector circuit is:

$$P_B = V_{CC}I_C = 16 \times 20 \times 10^{-3} = 320 \text{ mW}.$$

$$\text{Thus, } \eta = \frac{P_o}{P_B} \times 100\% = \frac{50}{320} \times 100\% = 15.6\%,$$

which is a little less than the 17.5% found for the CB amplifier in Chapter 10 for the same load and approximately the same output power. This is to be expected, since a CB has its characteristics occupying more of the area near the vertical collector current axis, allowing a smaller value of $V_{C,min}$.

11-5. PHASE RELATIONSHIPS IN A CE AMPLIFIER

Let us determine the current and voltage waveforms for the conditions given in Example 11-1: $R_L = 0.4$ K ohm, $V_{CC} = 16$ V, I_B bias $= 0.4$ mA, $I_s = 0.4$ mA p-p, $I_C = 20$ mA, $I_c = 20$ mA p-p.

$v_o = v_C = V_{CC} - i_C R_L$. With no v_s, $I_C = 20$ mA, $R_L = 0.4$ K ohm, $V_{CC} = 16$ V, $I_B = 0.4$ mA $\therefore v_o = 16 - 20 \text{ mA} \times 0.4 \text{ K ohm} = 16 - 8 = 8$ V.

If $i_C = 30$ mA, $i_B = 0.6$ mA,

$$v_o = 16 - 30 \text{ mA} \times 0.4 \text{ K ohm}$$
$$= 16 - 12 = 4 \text{ V}.$$

If $i_C = 10$ mA, $i_B = 0.2$ mA,

$$v_o = 16 - 10 \text{ mA} \times 0.4 \text{ K ohm}$$
$$= 16 - 4 = 12 \text{ V}.$$

The corresponding waveforms are shown in Fig. 11-6. Note that phase inversion occurs in the CE amplifier.

11-6. INPUT WAVEFORM CONSIDERATION

It is to be noted from Fig. 11-6(c) that, as in the CB amplifier, there is once again distortion of the voltage waveform at the input, with i_B sinusoidal. However, this will only occur if the source resistance R_s is large compared with the input resistance to the transistor, R_i. That is, the nonlinear relationship between i_B and v_{BE} in the input characteristic represents a changing input resistance over a full cycle of the input. If R_s is large, it "masks out" this small change in R_i, so that if v_s is sinusoidal, then the input current, $i_s = \frac{v_s}{R_s + R_i} \approx \frac{v_s}{R_s}$, and i_s is also sinusoidal.

However, this sinusoidal current flowing into a changing resistance R_i will develop a nonsinusoidal or distorted sine wave, v_{BE} as shown in Fig. 11-6(c). This is termed *current* drive, and in large-signal amplifiers will cause appreciable distortion in the output. The reason for this may be seen by referring to the ac load line drawn in Fig. 11-2(b).

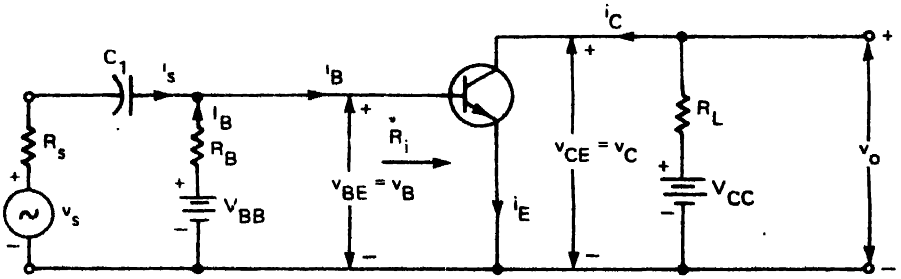


Fig. 11-5. Basic CE amplifier for examining waveforms.

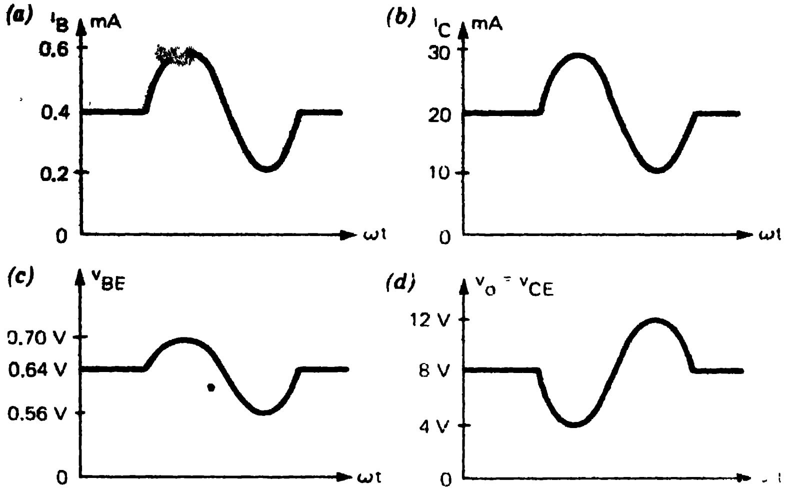


Fig. 11-6. Input and output waveforms in a CE amplifier.

Observe that a change in base current of 0.2 mA from R′ to T causes a larger change in output voltage v_{CE} than a change of 0.2 mA from P′ to S. This is due to the more insensitive region of the characteristic curves around P′ compared with R′. Put another way, the curves experience crowding for larger values of base and collector current, with the result that a base current drive that is symmetrical about the quiescent point will produce an unsymmetrical output in a large signal amplifier. The effect is really one of reduced forward current transfer ratio h_{fe} at large collector currents, as covered in Chapter 9.

Since the input resistance to the transistor is typically at least 1 K ohm, then the more usual case will be that R_s is smaller than R_i. This means that $i_s = \frac{v_s}{R_s + R_i} \approx \frac{v_s}{R_i}$, and i_s will be different from a sine wave (assuming v_s is a sinusoid), but the input voltage v_{BE} *will* be a sine wave. This is called *voltage* drive and in large-signal amplifiers can help reduce distortion if the load is properly selected. A typical waveform of base current resulting from voltage drive is shown in Fig. 11-7. It is evident that the greater positive swing from the quiescent value of I_B compared with the negative swing will *compensate* for the unequal spacing of the collector characteristics.

Thus, less distortion will occur in the output waveform if R_s is small, allowing v_{BE} to be sinusoidal but causing i_B to be distorted. So, once again, the effect of R_s cannot be ignored. If the input signal is small, however, as in Example 11-1, there will be little difference in output distortion, no matter whether R_s is large or small.

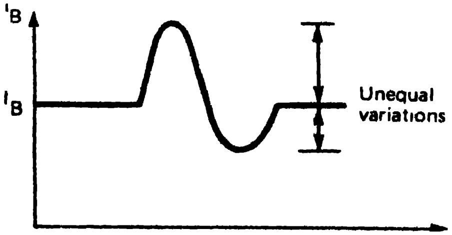


Fig. 11-7. CE amplifier base current resulting from voltage drive, $R_s << R_i$.

11-7. HARMONIC DISTORTION CALCULATION

To illustrate the methods of calculating harmonic distortion that were covered in a general manner in Chapter 7 and to see the differences that result from current and voltage drive, consider the following example.

Example 11-3. A 40319 power transistor is used as an audio driver stage to provide a signal to the output stage, which acts effectively as a load resistance R_L of 44 ohms, as shown in Fig. 11-8. Determine: (1) the percent second harmonic distortion in v_o using the 3-point method, and (2) the total harmonic distortion using the 5-point method for both of the following cases:

Current drive, signal current = 4 mA p-p.

Voltage drive, signal voltage = 0.30 V p-p.

Use the input and collector characteristics of the 40319 shown in Fig. 11-9, on which the 1-watt maximum dissipation hyperbola has been drawn for the collector at a 25° C ambient temperature.

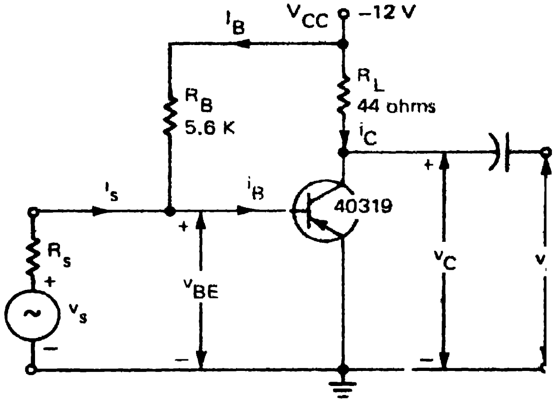


Fig. 11-8. CE audio driver amplifier to be analyzed for distortion.

Solution

(a) Load Line, Ac and Dc

The extremities of the load line are:

(1) $V_{CE} = V_{CC} = -12\text{ V}, I_C = 0$

(2) $V_{CE} = 0, I_C = \dfrac{V_{CC}}{R_L} = \dfrac{-12}{44} = -273\text{ mA.}$

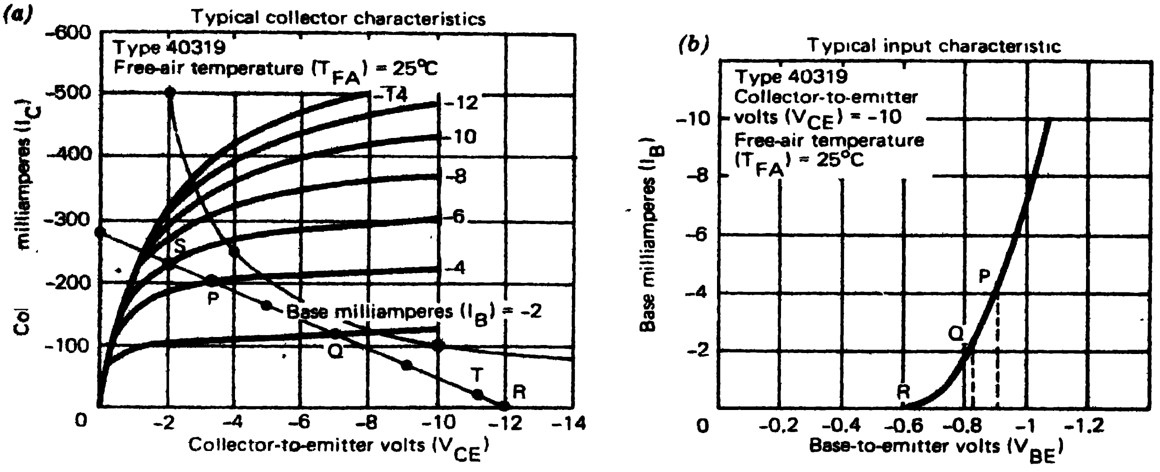


Fig. 11-9. Collector and input characteristics for PNP silicon 40319 transistor. (Courtesy of RCA Corp.)

Note: The negative signs indicate that with a PNP transistor the voltage polarities and current directions are opposite to those assumed in Fig. 11-8.

(b) Operating Point

For the given load line and maximum dissipation hyperbola, a reasonable bias would be at $I_B = 2$ mA, for which $V_{BE} \approx 0.8$ V. Thus,

$$I_B = \frac{V_{CC} - V_{BE}}{R_B} = \frac{12 - 0.8}{5.6 \times 10^3} = \frac{11.2 \text{ mA}}{5.6} = 2 \text{ mA}.$$

This determines Q for both current and voltage drive.

(c) Current Drive Harmonic Distortion

For a current drive of 4 mA p-p, which is understood to be symmetrical about $I_B = 2$ mA, the base current will vary from 0 mA at R to 4 mA at P. (Negative values for I_B on Fig. 11-9(a) indicate that the dc current flow is out from the base.) Reading off the values for current and voltage yields the following:

P: $I_{B,max} = 4$ mA, $V_{B,max} = 0.9$ V,

$I_{C,max} = 200$ mA, $V_{C,min} = 3.2$ V

Q: $I_B = 2$ mA, $V_{BE} = 0.82$ V,

$I_{C,Q} = 116$ mA, $V_C = 6.9$ V

R: $I_{B,min} = 0$ mA, $V_{B,min} = 0.60$ V,

$I_{C,min} = 0$ mA, $V_{C,max} = 12$ V

Observe that the increase in collector current from the quiescent value is only 84 mA, compared with a negative swing of 116 mA, and the input voltage waveform is flattened on its positive excursion, so that the waveforms of base current and voltage are similar to those in Fig. 11-6(a) and (c).

(1) Let us use the 3-point method to calculate the second harmonic distortion:

$$A_1 = \frac{I_{max} - I_{min}}{2} = \frac{200 - 0}{2} = 100 \text{ mA}$$

$$A_o = A_2 = \frac{I_{max} + I_{min} - 2I_Q}{4}$$

$$= \frac{200 + 0 - 2 \times 116}{4} = -8 \text{ mA}.$$

This indicates that the no-signal collector current of 116 mA will decrease to 108 mA upon application of the signal. A_1 is the amplitude of the desired fundamental, and $A_2 = 8$ mA is the amplitude of the second harmonic. Thus,

$$D_2 = \frac{|A_2|}{|A_1|} \times 100\% = \frac{8}{100} \times 100\% = 8\%$$

second harmonic distortion.

(2) To determine the total harmonic distortion using the 5-point method requires the determination of $I_{\frac{1}{2}}$ and $I_{-\frac{1}{2}}$. These have been designated as the collector currents corresponding to one half of the positive peak base current and one half of the negative peak base current. In this case, with a current drive of 4 mA p-p from a bias level of 2 mA, this corresponds to base currents of 3 mA and 1 mA respectively. From Fig. 11-9(a) we obtain for the collector currents, by estimating the 3 mA and 1 mA curves,

$I_{\frac{1}{2}} = 160$ mA and $I_{-\frac{1}{2}} = 60$ mA.

Substituting these values along with the max and min values of I_C in the following equations gives the required amplitudes.

$$A_o = \frac{1}{6}\left(I_{max} + 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} + I_{min}\right) - I_Q$$

$$= -9.4 \text{ mA}$$

$$A_1 = \frac{1}{3}\left(I_{max} + I_{\frac{1}{2}} - I_{-\frac{1}{2}} - I_{min}\right) = 100 \text{ mA}$$

$$A_2 = \frac{1}{4}\left(I_{max} - 2I_Q + I_{min}\right) = -8 \text{ mA}$$

$$A_3 = \frac{1}{6}\left(I_{max} - 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} - I_{min}\right) = 0$$

$$A_4 = \frac{1}{12}\left(I_{max} - 4I_{\frac{1}{2}} + 6I_Q - 4I_{-\frac{1}{2}} + I_{min}\right)$$

$$= -1.3 \text{ mA}$$

$$D_2 = \frac{|A_2|}{|A_1|} = \frac{8}{100} \times 100\% = 8\%$$

$$D_3 = \frac{|A_3|}{|A_1|} = \frac{0}{100} \times 100\% = 0\%$$

$$D_4 = \frac{|A_4|}{|A_1|} = \frac{1.3}{100} \times 100\% = 1.3\%$$

$$D = \sqrt{{D_2}^2 + {D_3}^2 + {D_4}^2 + \ldots}$$

$$\approx \sqrt{8^2 + 0^2 + 1.3^2} = 8.1\%.$$

Note that there is more shift in the operating point than was predicted by the 3-point method, which is less accurate for determining A_o.

Thus, the total harmonic distortion (due to the first three harmonics) will be in the order of 8.1% when using current drive of the magnitude of 4 mA p-p. Now let us see what the distortion is if voltage drive is employed.

(d) Voltage Drive Harmonic Distortion

At the Q point, $V_{BE} = 0.82$ V, so a voltage drive of 0.30 V p-p (0.15 V peak) will make $V_{B,max} = 0.97$ V and $V_{B,min} = 0.67$ V. Using these values on the input characteristic to determine the corresponding base current, we find that $I_{B,max} = 6$ mA while I_{B}, min = 0.2 mA from a bias value of $I_B = 2$ mA.

Evidently there is a much greater positive swing than the negative excursion. Let us find the corresponding values of collector current and determine the new harmonic distortion:

S: $V_{B,max} = 0.97$ V, $I_{B,max} = 6$ mA,

$I_{C,max} = 233$ mA, $V_{C,min} = 2$ V

Q: $V_{BE} = 0.82$ V, $I_B = 2$ mA, $I_C = 116$ mA,

$V_C = 6.9$ V

T: $V_{B,min} = 0.67$ V, $I_{B,min} = 0.2$ mA,

$I_{C,min} = 14$ mA, $V_{C,max} = 11.4$ V.

Observe that the positive swing in collector current is 117 mA, compared with 102 mA on the negative peak. Evidently voltage drive has had a compensating effect, so that the waveforms of i_C and V_{CE} will now be similar to those in Fig. 11-6(b) and (d), while the base current waveform will be similar to that in Fig. 11-7. The base-emitter voltage is, of course, sinusoidal.

(1) Using the 3-point method for second harmonic distortion calculation, we obtain:

$$A_1 = \frac{I_{max} - I_{min}}{2} = \frac{233 - 14}{2} = 109.5 \text{ mA}$$

$$A_o = A_2 = \frac{I_{max} + I_{min} - 2I_Q}{4}$$

$$= \frac{233 + 14 - 2 \times 116}{4} = 3.75 \text{ mA}.$$

In this case, the operating point will tend to shift upward slightly since A_o is positive.

$$D_2 = \frac{|A_2|}{|A_1|} = \frac{3.75}{109.5} \times 100\% = 3.4\%.$$

This is a significant reduction in second harmonic distortion from the 8% with current drive.

(2) For the 5-point method, we must again determine $I_{\frac{1}{2}}$ and $I_{-\frac{1}{2}}$, but we must be very careful to use the correct values of base current. Although the base current varies from 2 mA up to 6 mA and down to 0.2 mA, the values of base current to use are *not* 4 mA and 1.1 mA. Since the condition under examination is voltage drive we must consider one-half of the positive and negative peaks of the excitation voltage, then determine base current and finally collector current.

$$V_{B\frac{1}{2}} \quad 0.82 + \frac{0.97 - 0.82}{2}$$

$$= 0.82 + 0.075 = 0.895 \text{ V.}$$

$$V_{B-\frac{1}{2}} = 0.82 - 0.075 = 0.745 \text{ V.}$$

Corresponding values of base current and collector current are:

$I_{B\frac{1}{2}} = 3.9$ mA	$I_{\frac{1}{2}} = 180$ mA
$I_{B-\frac{1}{2}} = 1$ mA	$I_{-\frac{1}{2}} = 60$ mA

Making the appropriate substitutions in the equations used earlier, the amplitudes are as follows:

$A_o = 5.1$ mA, $A_1 = 113$ mA, $A_2 = 3.75$ mA,

$A_3 = -3.5$ mA, $A_4 = -1.4$ mA.

The percent harmonic distortion is:

$D_2 = 3.4\%$, $D_3 = 3.1\%$, $D_4 = 1.2\%$

and the total harmonic distortion, $D = 4.7\%$.

Note that this distortion of 4.7% with voltage drive compares with 8.1% with current drive, although in both cases the base-emitter signal voltage was 0.3 V p-p. Finally, let us determine the approximate values of R_s for current and voltage drive conditions.

$$\text{For current drive, } R_i = \frac{V_{be,p\text{-}p}}{I_{b,p\text{-}p}}$$

$$= \frac{0.30 \text{ V,p-p}}{4 \times 10^{-3} \text{ A,p-p}}$$

$$= 75 \text{ ohms.}$$

$$\text{For voltage drive, } R_i = \frac{0.30}{5.8 \times 10^{-3}} \approx 52 \text{ ohms.}$$

Thus, if $R_s = 600$ ohms for example, the conditions for current drive would be met (with its attendant higher harmonic distortion), while, if $R_s = 5$ ohms or less, voltage drive would result, with a reduction in harmonic distortion. However, it is possible with voltage drive, if R_s is very small, that *overcompensation* will occur and no appreciable decrease in total harmonic distortion will result.

11-8. OSCILLOSCOPE DISPLAY OF CHARACTERISTIC CURVES

The analysis carried out in Example 11-3 used the average characteristics of the transistor, but the 40319's h_{FE} may be anywhere from 35 to 200, so that the actual operation may be different from that predicted. Although many circuits have features to reduce the effect of variation in h_{FE} from one transistor to another. it is very useful if one can obtain the characteristic curves for a given transistor without tedious point by point plotting.

One inexpensive method is illustrated in Fig. 11-10. The operation of the circuit is as follows:

V is adjusted so that its peak value is somewhat lower than the maximum permissible value for V_{CE}. The diode D1 half-wave rectifies the 60 Hz input voltage, so that current will be allowed to flow through the transistor in one direction only, at a 60 Hz rate. The amount of current flowing in the collector will be determined by the base bias current I_B, set manually. The voltage across the transistor will deflect the oscilloscope trace horizontally, so that the horizontal input will

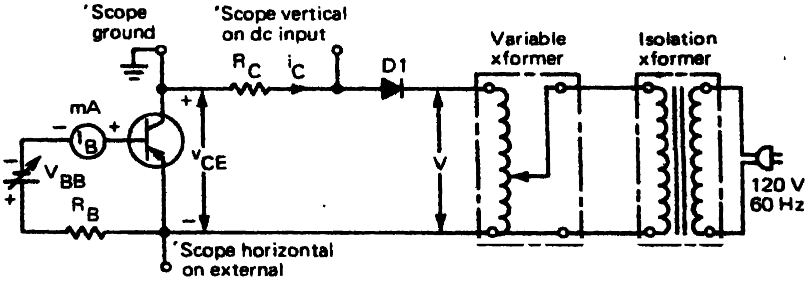


Fig. 11-10. Method of displaying transistor characteristic curves on oscilloscope.

have to be calibrated to give the actual voltage from collector to emitter.

The collector current i_C, flowing through R_C, will develop a voltage which will deflect the oscilloscope trace vertically, proportional to the current. Using the vertical calibration of the oscilloscope and the value of R_C will allow the determination of collector current per unit square of the oscilloscope screen.

The curves will appear one at a time for each value selected for I_B and must be drawn on graph paper to provide the whole set of curves. These curves may in some cases be in a different quadrant from the one in which they should be drawn, depending upon the particular oscilloscope used.

By proper selection of V and R_C, the maximum dissipation of the transistor will not be exceeded.

If in some way the base current could be stepped through its values (of 0, 0.1 mA, 0.2 mA, etc.) at some rapid rate, automatically repeating, then the whole family of collector characteristics would be displayed simultaneously and could be photographed. This is the method used in commercial curve tracers, one of which is shown in Fig. 11-11. In this model, horizontal and vertical calibrations are provided in a readout module, which also shows the value of base current per step and automatically displays the value of β per division. NPN or PNP transistors may be tested up to 220 watts of collector power and up to 1500 V. Diodes and field effect transistors may also be examined. An optional attachment allows high current pulses of 200 A peak in the collector to observe the characteristics of high power transistors and diodes.

11-9. COMPARISON OF BASIC TRANSISTOR AMPLIFIERS

Although the common-collector (or emitter follower) will not be covered in detail until Chapter 17, we can state for comparison purposes that it has a voltage gain less than one, a high current gain, high input resistance and low output resistance. A common-emitter amplifier has a medium input and medium output resistance when compared with a CB or CC circuit. However, it has both current gain and voltage gain, with the greatest power gain of all three. The output voltage waveform is inverted with respect to the input. Fig. 11-12 summarizes the three transistor configurations, with both PNP and NPN transistors shown, along with their comparative gains and resistances.

11-10. SUMMARY

(1) The analysis of a class A CE amplifier involves drawing the dc load line, locating the operating point at $I_B \approx \frac{V_{BB}}{R_B}$, drawing

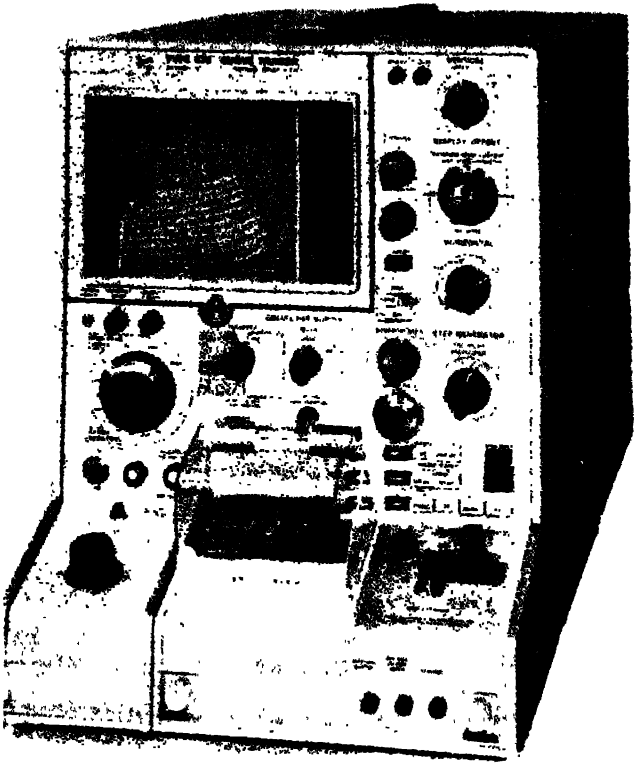


Fig. 11-11. Commercial curve tracer with parameter readout. (Courtesy of Tektronix, Inc.)

the ac load line through the operating point with a slope given by $\frac{-1}{R_e}$ where $R_e = R_o \| R_L$, and reading off the corresponding values of collector current and voltage from the ac load line resulting from a specified base current drive.

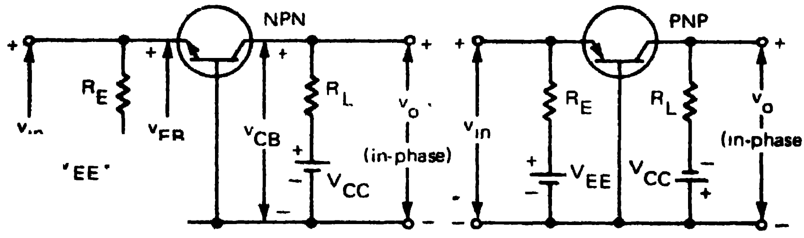
(2) One of the greatest advantages of a CE amplifier is the ability to operate from a single supply voltage, using a resistor from V_{CC} to the base to provide bias current.

(3) A CE amplifier enjoys both a voltage and a current gain, with the result that its power gain is usually larger than in either a CB or CC.

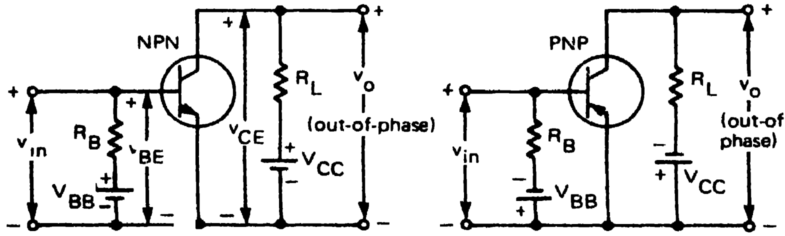
(4) The input resistance to a CE amplifier is in the order of 1 K ohm and is thus much higher than in a CB, but lower than a CC. Its output is inverted with respect to its input voltage.

(5) The maximum theoretical conversion efficiency of a class A amplifier with a

Common Base $A_v > 1, A_i < 1$. Low R_{in}, high R_{out}.


NPN
PNP
+
−
R_E
R_L
v_{in}
V_{EB}
V_{EE}
V_{CB}
V_{CC}
v_o
(in-phase)

Common Emitter $A_v > 1, A_i > 1, A_p$ very high. Med. R_{in}, med. R_{out}.


NPN
PNP
+
−
R_B
R_L
v_{in}
V_{BE}
V_{BB}
V_{CE}
V_{CC}
v_o
(out-of-phase)

Common Collector $A_v < 1, A_i > 1$. High R_{in}, very low R_{out}.

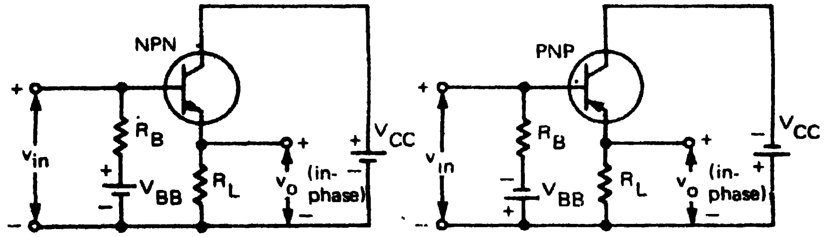

NPN
PNP
+
−
R_B
R_L
v_{in}
V_{BB}
V_{CC}
v_o
(in-phase)

Fig. 11-12. Basic PNP and NPN transistor amplifiers.

direct-coupled resistive load is 25%.

(6) If the source resistance R_s is large compared with a CE's input resistance R_i, a condition of current drive results such that base current variation is symmetrical but base-emitter voltage is not. For large signal inputs this can cause considerable distortion in the output.

(7) Voltage drive results when $R_s << R_i$, so that the unequal variations in base current resulting from large-signal operation tend to compensate for the insensitive region of the characteristic curves at higher collector current, keeping harmonic distortion lower.

(8) A transistor's characteristic curves may be displayed on an oscilloscope using a manual control of base current or by using a commercial curve tracer that permits a photograph of the complete set of curves.

(9) In normal amplifier operation, the transistor's base-emitter junction is always forward biased but the collector-base junction is reverse biased.

QUESTIONS

Q11-1. What is the main advantage of a CE over a CB as far as biasing is concerned?

Q11-2. Why is loading of the input signal by the bias resistor not generally a problem in a CE amplifier compared with a CB?

Q11-3. What is an approximate expression for the base bias current in a CE amplifier? Would this expression be more accurate for a germanium transistor than for a silicon? Why?

Q11-4. If a CE amplifier's current gain is 50, what can you state regarding its h_{fe}?

Q11-5. An instructor remarked that a CE's input resistance is usually at least *one order of magnitude* larger than that for a CB. What is meant by this statement? Give values to explain.

Q11-6. A CE provides more *isolation* between input and output than does a CB. What is meant by this?

Q11-7. What is the maximum theoretical conversion efficiency of a directly-coupled class A amplifier? Does an amplifier's actual conversion efficiency have anything to do with the drain on the power supply? Explain.

Q11-8. Why is it theoretically possible to obtain a larger conversion efficiency in a CB amplifier than in a CE? Is the maximum still 25%?

Q11-9. What type of Lissajous figure would you expect on an oscilloscope when checking the phase shift between input and output in a CE amplifier in the mid-frequency range? How would this change at higher frequencies?

Q11-10. If the input voltage between base and emitter is sinusoidal when operating a CE as a large-signal amplifier, what would you expect the waveform of base current to be like? What type of signal drive is this called? What advantages does it have?

Q11-11. What conditions are required to provide current drive in a large-signal CE amplifier? What shape is the voltage waveform from base to emitter? What causes this distortion?

Q11-12. Under what conditions will voltage and current drive not cause any difference in distortion in the output?

Q11-13. When might voltage drive cause more total harmonic distortion in the output than when current drive is used?

Q11-14. Sketch a circuit to display the characteristic curves for an NPN transistor on an oscilloscope.

Q11-15. It was noted that a CE amplifier could be operated from a single power supply. Can a common collector? Devise a voltage divider method so that a single battery can provide proper bias for a CB amplifier. (Hint: Use a resistor in the base lead bypassed by a capacitor.)

PROBLEMS

P11-1. A 40421 germanium PNP transistor, whose characteristics are given on the following page, is used in an audio amplifier circuit like that in Fig. 11-3. $V_{CC} = 30$ V, $R_L = 6$ ohms.

(a) Draw the dc (and ac) load line.

(b) Determine R_B to provide a base bias of 20 mA.

(c) Specify the Q point: V_{CE}, I_C, and V_{BE}.

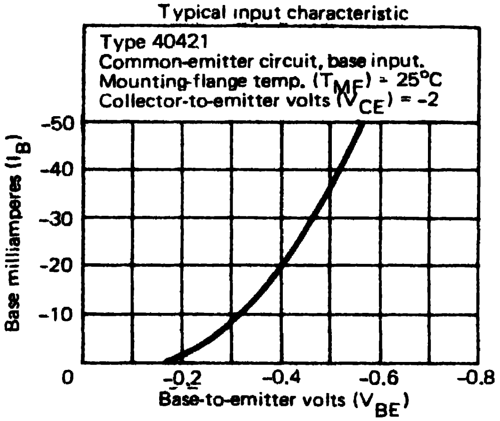
(d) Determine the peak-to-peak collector current for a 30 mA p-p input base current, and hence determine the current gain.

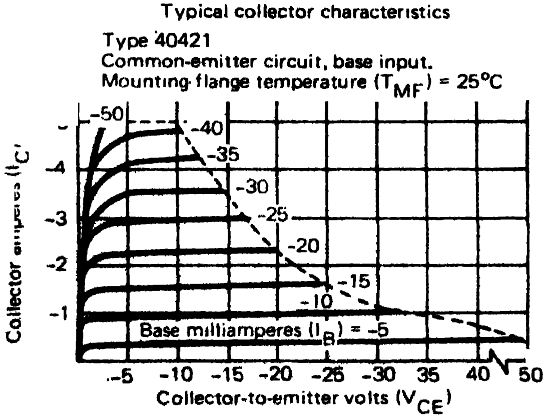
(e) Determine the peak-to-peak output voltage for a 30 mA p-p input.

(f) Obtain the base-emitter peak-to-peak voltage for (e) and obtain the voltage gain.

(g) Determine the power gain as a ratio and in decibels.

(h) Determine the ac input resistance from (f). Do you think R_B will cause any appreciable loading?


Typical input characteristic


Typical collector characteristics

Characteristics of the 40421 transistor for P11-1. (Courtesy of RCA Corp.)

(i) Current drive was used in this problem. What value of R_s would you suggest to ensure current drive? Would voltage drive be beneficial for this load line?

P11-2. If an additional 6-ohm load, R_o, is connected across the output of the amplifier in P11-1 (after C_2), draw the ac load line and repeat P11-1(d) through (g).

P11-3. For the amplifier in P11-1, assume a 40 mA p-p base input signal is applied instead of 30 mA p-p. Calculate:

(a) The dc input power supplied by the battery to the collector circuit.

(b) The ac output power delivered to R_L.

(c) The conversion efficiency of the circuit.

P11-4. For the conditions in P11-3 (40 mA p-p input signal, R_L = 6 ohms, V_{CC} = 30 V), draw the waveforms of i_B, i_C, v_{BE}, and v_{CE} as in Fig. 11-6. Show numerical values.

P11-5. (a) Using the base-emitter bias voltage from P11-1(c) and the ac base-emitter voltage from P11-1(f) with the corresponding collector currents, calculate the percent second harmonic distortion using the 3-point method. This is a condition of current drive, with unsymmetrical voltage swings.

(b) Use the same base-emitter bias voltage as in (a) and the same peak-to-peak base-emitter voltage, but now assume voltage drive, so that the base current swing is unsymmetrical. Recalculate the second harmonic distortion. Do you see how voltage drive is not beneficial if the collector characteristics are equally spaced, as they are for this transistor?

P11-6. Sketch the waveforms of base current and voltage for the two conditions in P11-5.

P11-7. A silicon NPN 2N3054 transistor is used in a CE amplifier circuit as in Fig. 11-8. The characteristic curves are shown

below. $V_{CC} = 40$ V, $R_L = 16$ ohms. The amplifier is to be analyzed for second harmonic distortion under the conditions of current and voltage drive.

(a) Draw the dc (and ac) load line.

(b) Determine the value of R_B to provide a base bias current of 40 mA.

(c) Specify the operating point: V_{CE}, I_C, V_{BE}.

(d) Calculate the percent second harmonic distortion assuming:

(1) Current drive with a 60 mA p-p input signal.

(2) Voltage drive with a 0.5 V p-p base-emitter voltage.

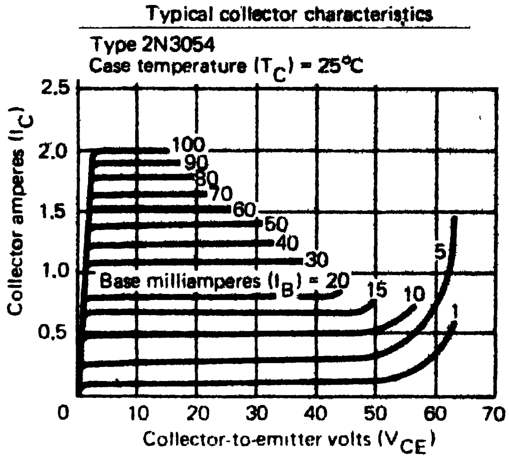
(e) Sketch the waveforms of base current and voltage for the above two drives.

P11-8. Repeat P11-7(d) and (e) using:

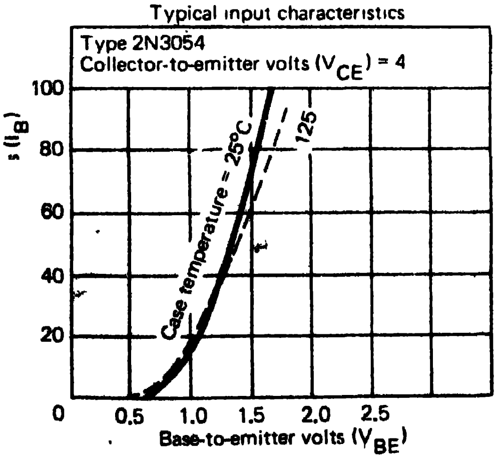
(a) Current drive with an 80 mA p-p input signal.

(b) Voltage drive with a 0.90 V p-p base-emitter voltage.

P11-9. Calculate the total harmonic distortion for the two conditions in P11-8.



Characteristic curves of the 2N3054 transistor. (Courtesy of RCA Corp.)



12

THERMAL STABILITY I: TRANSISTOR BIASING

12-0. INTRODUCTION

A transistor in a common-emitter amplifier can be biased in many different ways, but each has a different effect on the thermal stability of the operating point. Once an operating point has been chosen, the problem is how can it be maintained in the face of temperature increase, so that distortion or thermal runaway will not occur. The stability of different circuits will be examined, and methods of compensation will be discussed.

12-1. FACTORS CONTRIBUTING TO THERMAL INSTABILITY

It was shown that the collector current for a common-emitter amplifier is given by the equation:

$$I_C = \beta I_B + (\beta + 1) I_{CO}. \qquad (12\text{-}1)$$

Note: $\beta = h_{FE}, I_{CO} = I_{CBO}$.

This equation contains three variables, all dependent upon temperature.

(a) It has already been shown that β is strongly dependent upon temperature, no matter whether germanium or silicon is being considered, as shown by Fig. 12-1(a).
(b) Fig. 12-1(b) indicates how the base current, I_B, increases with temperature if V_{BE} is held constant, which represents a decrease in dc input resistance. The effect is the same for both germanium and silicon, the curves shifting to the left at the rate of 2.5 mV/°C for constant I_C. In silicon transistor amplifiers, this effect is dominant over the increase in leakage current I_{CO}.
(c) The increase in leakage current has been indicated as a doubling of I_{CO} for every 10°C increase in temperature, for either silicon or germanium. However, the fact that I_{CO} is very much smaller to start with in silicon than it is with germanium makes a silicon transistor amplifier much more immune to temperature change *due to this effect.*

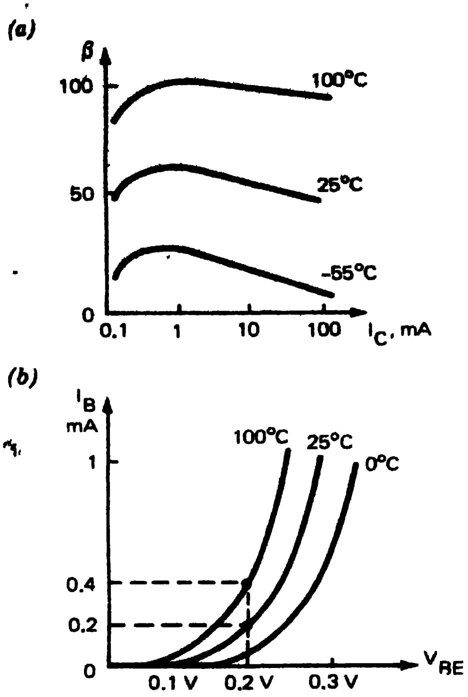


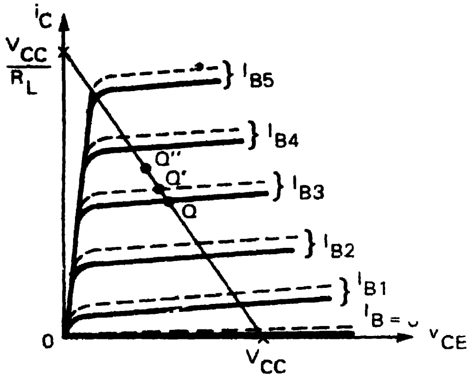
Fig. 12-1. (a) Variation of β with temperature for typical germanium transistor. (b) Input characteristics for typical germanium transistor in CE circuit at various temperatures.

12-2. EFFECT OF TEMPERATURE INCREASE

The net result of these increases in β, I_B and I_{CO} should be evident from the equation:

$$I_C = \beta I_B + (\beta + 1)I_{CO}.$$

I_C will increase, shifting the characteristic curves upward and changing the operating point, as shown in Fig. 12-2.



Q @T_1.

Q' @$T_2 > T_1$; same bias I_B as at T_1.

Q'' @T_2 with bias I_B increased due to temp. $T_2 > T_1$ (effect of decreased input resistance as in (b)).

Fig. 12-2. Showing a change in the operating point in a CE amplifier due to temperature change. The solid lines are at a temperature T_1 and the dotted lines at a temperature $T_2 > T_1$ or for a transistor at T_1 with a higher value of β.

The movement of the operating point toward the saturation portion of the characteristics could cause clipping and bad distortion of the output if the magnitude of the input signal were unchanged. What may be an acceptable bias point at one temperature may become very poor at an increased (or decreased) temperature.

It should be noted that the shift in the characteristic curves shown above may come about due to a different transistor being used in the amplifier, and not by a change in temperature. A 2N109 for example has a β, as stated in the transistor manual, between 50 and 150. This 3:1 spread is not unusual and can itself change the operating point. It is evident that keeping I_B constant will not compensate for changes in β. If β *increases* (either due to a new transistor being used or an increase in temperature taking place), then I_B must *decrease* to keep I_C constant. This method is used also to compensate for increases in I_{CO} due to temperature.

Consider the equation:

$$I_C = \beta I_B + (\beta + 1)I_{CO}.$$

If I_{CO} increases, I_C will increase by $(\beta + 1)$ times as much. The increased I_C will increase the temperature of the transistor collector junction, which in turn will cause I_{CO} to increase. It is possible for this effect to be cumulative, amounting to "thermal runaway" in which the transistor burns out at the worst or at least causes bad distortion of the output.

If, however, I_B is made to decrease automatically with temperature increase (by circuit modification), then the decrease in βI_B will compensate for the increase in $(\beta + 1)I_{CO}$, keeping I_C nearly constant.

12-3. STABILITY FACTOR S

The extent to which the stabilizing of I_C is successful (in the face of variations in I_{CO}) is measured by a stability factor S.

This is defined as the rate of change of collector current I_C with respect to the leakage current I_{CO}, holding both β and I_B (V_{BE}) constant.

$$S = \frac{\partial I_C}{\partial I_{CO}} \simeq \frac{dI_C}{dI_{CO}} \simeq \frac{\Delta I_C}{\Delta I_{CO}} \text{ ; } \beta, I_B \text{ constant.} \quad (12\text{-}2)$$

(This is the derivative of I_C with respect to I_{CO}.)

Consider the CE amplifier using fixed bias shown in Fig. 12-3.

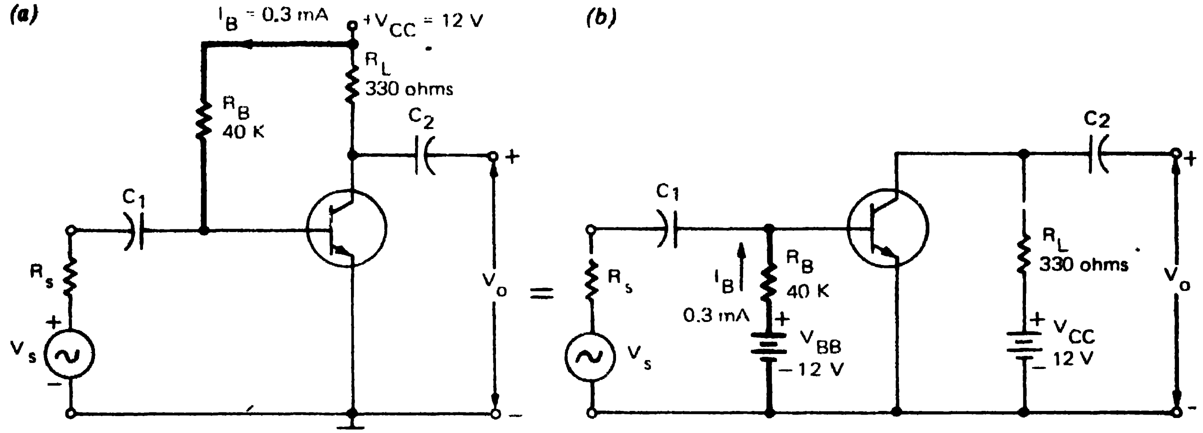


Fig. 12-3. (a) CE amplifier with fixed bias. (b) Alternative way of drawing the CE amplifier showing V_{BB} as a separate supply.

Since $I_C = \beta I_B + (\beta + 1) I_{CO}$ then

$$S = \frac{dI_C}{dI_{CO}} = \beta + 1. \qquad (12\text{-}3)$$

(By differential calculus)

If $\beta = 100$, then $S = 101$, which means I_C changes 101 times as much as any change in I_{CO}. This says that I_C is very dependent upon I_{CO} and hence upon temperature. This is true no matter whether a silicon or germanium transistor is being used, although in a silicon it may not be as troublesome, depending upon the size of the transistor and its application, since I_{CO} is as little as one-thousandth the leakage of a germanium transistor.

However, a silicon transistor is just as dependent upon temperature as a germanium as far as the increase in I_B is concerned and must be stabilized from this standpoint. Since the stabilizing of a transistor against variations in I_{CO} also stabilizes it adequately as far as changes in β and I_B are concerned (which can be very troublesome in a silicon transistor) it is important to check the stability figure, S, of a transistor amplifier, no matter whether it is germanium or silicon.

The fact that S is very large in a fixed-bias circuit indicates poor thermal stability. It is our object to examine ways to reduce S—a smaller value for S indicating a greater stability, with S = 1 being the smallest and most stable possible.

12-4. COMMON-BASE STABILITY

Consider for a moment the equation for I_C in a common-base circuit:

$$I_C = \alpha I_E + I_{CO}. \qquad (12\text{-}4)$$

$$\text{Then } S \simeq \frac{dI_C}{dI_{CO}} = 1. \qquad (12\text{-}5)$$

Hence, the common-base circuit is not in need of bias stabilization.

12-5. COLLECTOR-TO-BASE BIAS

One of the ways to automatically reduce I_B with increase in temperature is to obtain the bias voltage from the collector of the transistor instead of from the collector supply voltage. This is illustrated in Fig. 12-4.

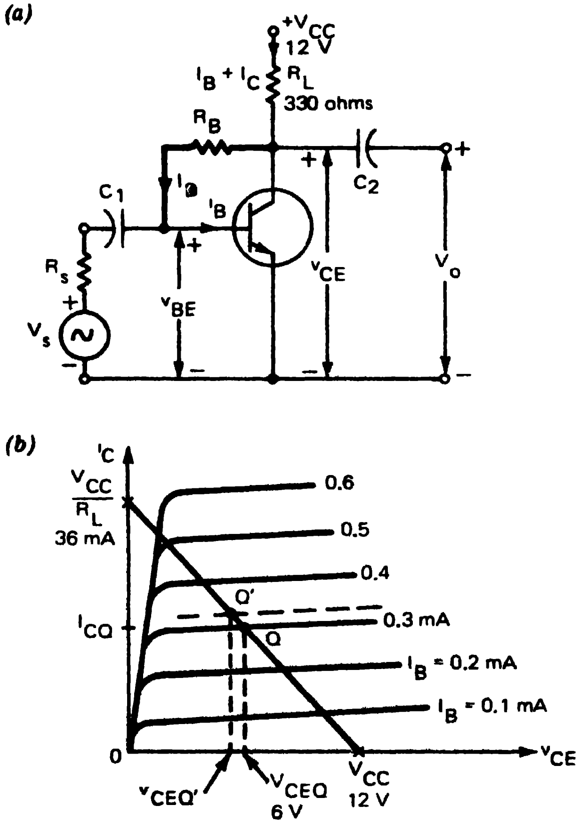


Fig. 12-4. (a) CE amplifier with collector-to-base bias. (b) Showing the shift in Q point due to an increase in temperature.

If the collector current tends to increase due to an increase in temperature or because the transistor has been replaced by one with a higher β, the greater dc voltage drop across R_L will result in a lower average value from collector to emitter, V_{CEQ}', with the result that $I_B \approx \frac{V_{CEQ}'}{R_B}$ will also decrease. This will tend to compensate for the original increase in I_C. Note, however, that it takes an increase in I_C to provide the decrease in I_B, so that the two will never compensate exactly.

The new stability factor S for this circuit can be shown to be:

$$S = \frac{\beta + 1}{1 + \frac{\beta R_L}{R_L + R_B}}, \qquad (12\text{-}6)$$

which is obviously smaller than $\beta + 1$ for the fixed bias circuit.

Example 12-1. Determine the bias resistors R_B, for fixed bias in Fig. 12-3(a) and collector-to-base bias in Fig. 12-4(a), and compare the stability factors S for the two circuits. Given: $V_{CC} = 12$ V, $R_L = 330$ ohms, $I_B = 0.3$ mA, $\beta = 100$, $V_{CEQ} = 6$ V (in Fig. 12-4(b)).

Solution. For the fixed bias circuit,

$$R_B \simeq \frac{V_{CC}}{I_B} = \frac{12 \text{ V}}{0.3 \text{ mA}} = 40 \text{ K ohms.}$$

For the collector-to-base bias circuit,

$$R_B \simeq \frac{V_{CEQ}}{I_B} = \frac{6 \text{ V}}{0.3 \text{ mA}} = 20 \text{ K ohms.}$$

The stability factor S for fixed bias = $\beta + 1$
100 + 1 = 101.

For collector-to-base bias:

$$S = \frac{\beta + 1}{1 + \frac{\beta R_L}{R_L + R_B}} = \frac{100 + 1}{1 + \frac{100 \times 330}{330 + 20{,}000}}$$

$$= \frac{101}{1 + 1.65} = 38.2.$$

This will make a considerably more stable circuit as far as variations in the Q point are concerned due to temperature causing I_{CO} to change.

However, how effective is this bias circuit in resisting changes in the operating point due to β variation arising from the use of different transistors?

It can be shown that if:

$$\beta R_L >> R_B, \tag{12-7}$$

then I_C is relatively insensitive to changes in β. Checking for the second case (collector-to-base bias) in Example 12-1,

$$\beta R_L = 100 \times 330 = 33 \text{ K ohms}$$

$$R_B = 20 \text{ K ohms.}$$

Thus βR_L is not quite twice R_B, so that the circuit is not completely insensitive to β variations. (A factor of 10 would be good, but is not always practical.)

However, compared with fixed bias, where $R_B = 40$ K ohms, the second method of biasing is twice as independent of β as is the first circuit, where $\beta R_L = 33$ K ohms < 40 K ohms.

12-6. DISADVANTAGES OF COLLECTOR-TO-BASE BIAS

There are two drawbacks associated with collector-to-base bias. First, since R_B has been decreased from 40 K ohms to 20 K ohms, a change in temperature which would increase I_B at a given V_{BE} would be much more successful with $R_B = 20$ K ohms than when $R_B = 40$ K ohms. That is, if the input resistance to the transistor decreases with an increase in temperature, the resulting increase in base current will be larger if 20 K ohms is in series with the applied bias voltage than if 40 K ohms is in series. This *increase* in I_B will tend to *offset* the required *decrease* in I_B brought about by collector-to-base biasing. In a silicon transistor circuit, this will be particularly true, and not much *overall* increase in stability may result.

The second drawback can be partially overcome. The fact that R_B is connected to the collector means that the ac signal present at the output will be fed back out of phase with the input; this is called negative feedback. Thus, the input will be reduced, decreasing the output. If R_B is split into two the signal can be taken to ground to reduce this effect by means of a capacitor, small in reactance compared with $\frac{R_B}{2}$, as shown in Fig. 12-5.

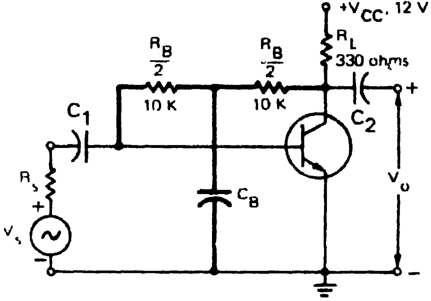


Fig. 12-5. Collector-to-base bias with capacitor C_B to reduce negative feedback.

This still provides dc bias stabilization. However, the result is a decrease in both the input and output ac impedances, since a 10 K ohm resistor has effectively been placed in parallel with the base-emitter junction and a 10 K ohm resistor from collector to emitter.

12-7. EMITTER BIAS

Consider the stability factor S of the collector-base bias circuit:

$$S = \frac{\beta + 1}{1 + \dfrac{\beta R_L}{R_L + R_B}}.$$

If R_L is very small, as in a transformer-coupled output load, S is not reduced very much. For example, say $R_L = 10$ ohms (dc resistance of primary of output transformer) and $R_B = 20$ K ohms.

Then,

$$S = \frac{100 + 1}{1 + \dfrac{100 \times 10}{10 + 20{,}000}} = \frac{101}{1 + \dfrac{1000}{20{,}000}}$$

$$= \frac{101}{1.05} = 96.$$

If it can be arranged so that, when the collector current increases, the forward bias across the base-emitter junction is made to decrease automatically, then I_B will be reduced and will serve to stabilize the dc operating point.

This is accomplished by placing a resistor in the emitter lead, as shown in Fig. 12-6. Since I_B is small then $I_E \approx I_C$.

Let us obtain an expression for the base bias current, with $V_s = 0$. Writing an equation around the input loop (see Fig. 12-7):

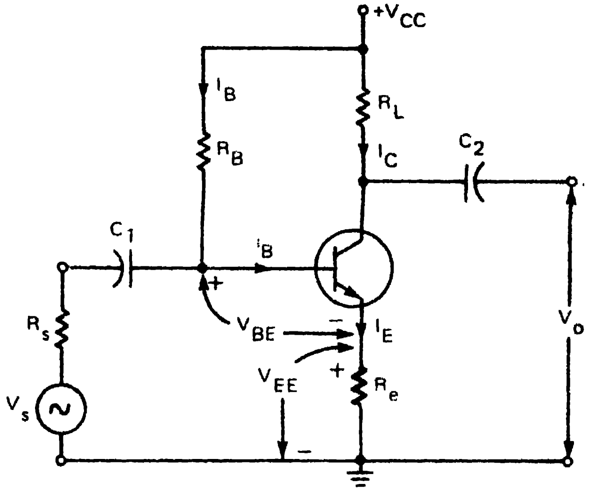


Fig. 12-6. Emitter bias due to R_e in the emitter lead.

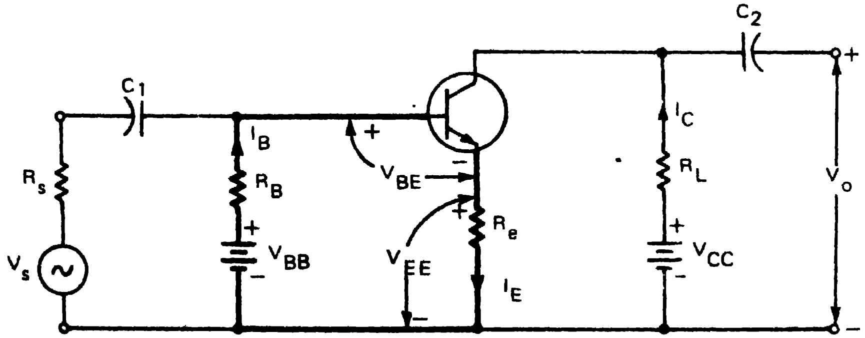


Fig. 12-7. Emitter bias circuit redrawn for analysis of the input loop.

$$V_{BB} = I_B R_B + V_{BE} + V_{EE} \qquad (12\text{-}8)$$

$$= I_B R_B + V_{BE} + I_E R_e. \qquad (12\text{-}9)$$

If $I_E \approx I_C$, then:

$$V_{BB} = I_B R_B + V_{BE} + I_C R_e. \qquad (12\text{-}10)$$

$$\therefore I_B R_B = V_{BB} - V_{BE} - I_C R_e. \qquad (12\text{-}11)$$

If $V_{BE} << V_{BB}$, then:

$$I_B \approx \frac{V_{BB} - I_C R_e}{R_B}. \qquad (12\text{-}12)$$

For a constant V_{BB} (may be equal to V_{CC}), an increase in I_C (for any reason) will reduce I_B, serving to offset the increase in I_C.

The stability factor for this new circuit is:

$$S = \frac{\beta + 1}{1 + \dfrac{\beta R_e}{R_e + R_B}}$$

$$= (\beta + 1) \times \frac{\left(1 + \dfrac{R_B}{R_e}\right)}{\beta + 1 + \dfrac{R_B}{R_e}}. \qquad (12\text{-}13)$$

Note that from the second form, if $\frac{R_B}{R_e}$ is very small, $\frac{R_B}{R_e} << 1$,

$$S = \beta + 1 \times \frac{1 + 0}{\beta + 1 + 0} = \frac{\beta + 1}{\beta + 1} = 1$$ (good thermal stability), and if $\frac{R_B}{R_e}$ is very large,

$$S = (\beta + 1) \times \frac{\dfrac{R_B}{R_e}}{\dfrac{R_B}{R_e}} = \beta + 1$$ (poor thermal stability). To make $\frac{R_B}{R_e}$ small, either R_e should be large, or R_B small, or both. However, the choice of R_B is somewhat restricted by the required bias current I_B. (A circuit to get around this restriction will be considered later.)

To get some idea of the stabilizing effect of R_e, consider adding a 100-ohm resistor in the emitter lead as in the following example.

Example 12-2. Determine the bias resistor R_B in Fig. 12-8(a) to provide a base bias current of 0.3 mA, using the characteristic curves in Fig. 12-8(b). Also calculate the stability factor S, comparing it with the fixed bias circuit of Fig. 12-3(a). $\beta = 100$.

Solution. Draw the dc load line:

(1) $i_C \quad 0, v_{CE} = V_{CC} = 12$ V.

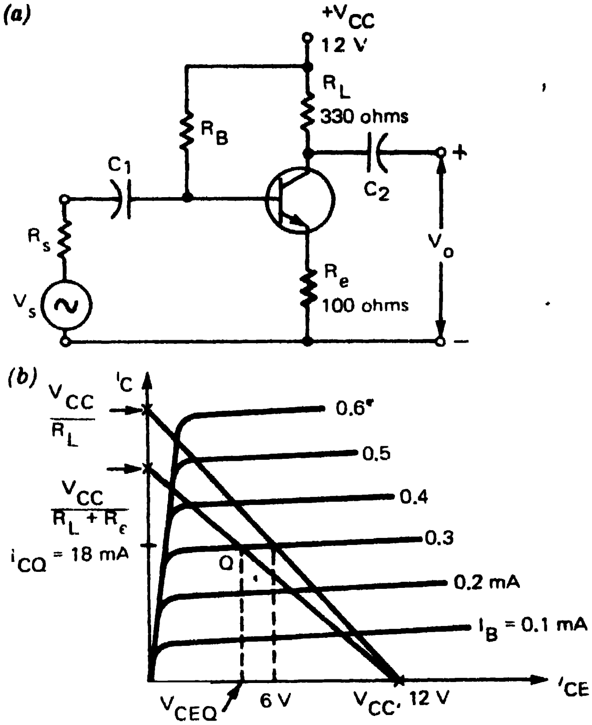


Fig. 12-8. (a) CE amplifier with emitter bias resistor, R_e, for Example 12-2. (b) Transistor characteristic curves with dc load lines.

(2) $v_{CE} = 0, i_C = \dfrac{V_{CC}}{R_L + R_e}$

$$= \frac{12}{0.33\text{ K} + 0.1\text{ K}} = \frac{12\text{ V}}{0.43\text{ K}}$$

$$= 28\text{ mA}.$$

For $I_B = 0.3$ mA, at Q: $I_{CQ} = 18$ mA, $V_{CEQ} = 4.25$ V. Note: Since the dc load line is not as steep now, a bias of 0.3 mA no longer provides a Q point in the center of the characteristics, and a lower value would be better, say 0.25 mA or 0.2 mA. This is one of the drawbacks of emitter bias. To operate at the same quiescent point, the value of V_{CC} would have to be increased if R_B is kept constant, or

R_B increased if V_{CC} held constant. Either way, a drop in efficiency is the price one pays for the increased stability.

It was shown before that:

$$I_B R_B \quad V_{BB} - V_{BE} - I_C R_e .$$

For a germanium transistor $V_{BE} \approx 0.2$ V; for Si, $V_{BE} \approx 0.6$ V. Assuming germanium:

$$R_B \quad \frac{V_{BB} - V_{BE} - I_C R_e}{I_B}$$

$$= \frac{12 - 0.2 - 18 \times 10^{-3} \times 0.1 \times 10^3}{0.3 \times 10^{-3}}$$

$$= \frac{12 - 0.2 - 1.8}{0.3 \times 10^{-3}} = \frac{10 \times 10^3}{0.3} = 33 \text{ K ohms.}$$

Calculating the stability factor:

$$S = \frac{\beta + 1}{1 + \dfrac{\beta R_e}{R_e + R_B}} = \frac{100 + 1}{1 + \dfrac{100 \times 100}{100 + 33{,}000}}$$

$$\frac{101}{1 + \dfrac{10 \text{ K}}{33.1 \text{ K}}} = \frac{101}{1 + 0.3} = \frac{101}{1.3} = 77.$$

This is not a great reduction in S, from 101 to 77, but this is due to the relatively large ratio of R_B/R_e:

$$\frac{R_B}{R_e} = \frac{33 \text{ K}}{0.1 \text{ K}} = 330.$$

Fig. 12-9 indicates the dependence of S upon R_B/R_e. The curves are drawn for different values of β, since the worst case of $S = \beta + 1$ is when R_B/R_e is very large.

The curves confirm that for a β of 100 there is little reduction in S when R_B/R_e is as much as 330.

To increase the stability of the above circuit would require the increasing of R_e to 200 or 300 ohms. But this would decrease the efficiency tremendously. The fact is that R_B is too large to decrease the ratio R_B/R_e simply by increasing R_e. The question is, how can the same bias current of $I_B = 0.3$ mA be supplied and yet have a reduction in R_B with the same value of V_{CC}?

Before answering this question let us see why the reduction of S is dependent upon a low value of R_B.

When the collector current increase reduces the forward bias from base to emitter it effectively *increases* the dc input resistance. This will only decrease I_B if this change in input resistance is comparable to the total resistance in series (as far as bias current is concerned) with the base, namely R_B. If R_B is 40 K ohms for example, a change in input resistance of a few ohms will not materially change I_B, and the effect of R_e on I_B is small.

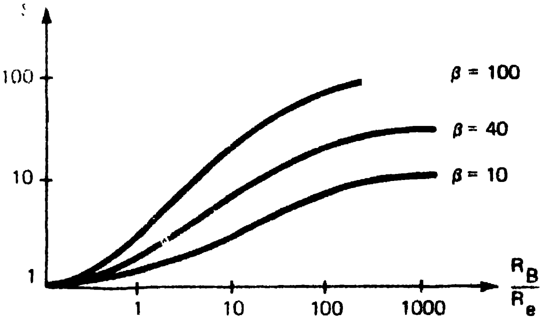


Fig. 12-9. Curves showing how the stability factor S depends upon R_B/R_e for various values of β.

12-8. VOLTAGE DIVIDER BIAS WITH EMITTER BIAS

In the example above, R_B had to be 33 K ohms with $V_{BB} = V_{CC} = 12$ V. If in some way V_{BB}

were reduced, R_B also could be reduced. Consider the voltage divider circuit in Fig. 12-10.

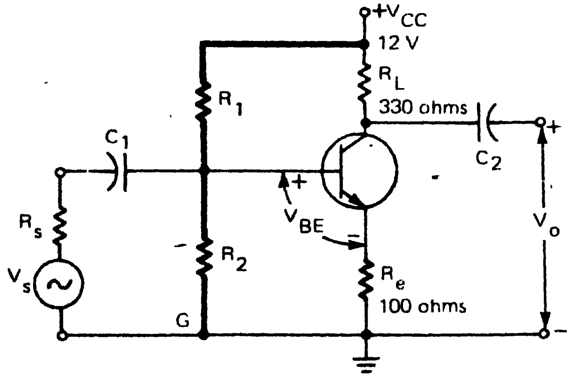

+V_{CC}
12 V
R_L
330 ohms
R_1
C_1
C_2
+
V_{BE}
−
R_s
V_s
R_2
G
R_e
100 ohms
V_o
+
−

Fig. 12-10. CE amplifier with voltage divider bias due to R_1 and R_2, and emitter bias due to R_e.

Consider only the bias circuit R_1, R_2, V_{CC}, as shown in Fig. 12-11.

$$R_B' = \frac{R_1 R_2}{R_1 + R_2} \qquad V_{BB} = V_{CC} \times \frac{R_2}{R_1 + R_2}$$

and (12-14)

$$S = \frac{\beta + 1}{1 + \dfrac{\beta R_e}{R_e + R_B'}}$$

$$= (\beta + 1) \times \frac{\dfrac{R_B'}{R_e}}{\beta + 1 + \dfrac{R_B'}{R_e}} . \qquad (12\text{-}15)$$

Obviously, V_{BB} will be less than V_{CC}, so that R_B' can be less (than R_B), and so will R_B'/R_e, reducing S further.

The selection of R_1 and R_2 for a required bias current is not an easy matter and is best illustrated by a numerical example.

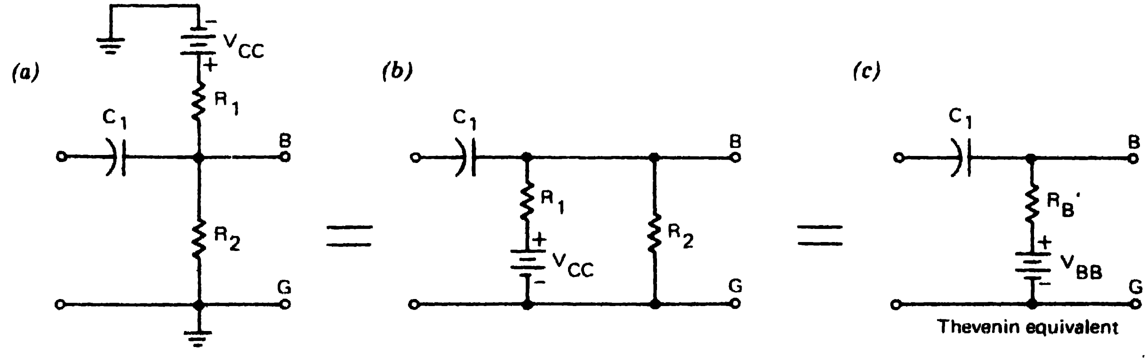

(a)
V_{CC}
+
−
C_1
R_1
B
R_2
G
(b)
C_1
R_1
+
−
V_{CC}
R_2
B
G
(c)
C_1
B
R_B'
+
−
V_{BB}
G
Thevenin equivalent

Fig. 12-11. (a) Voltage divider bias due to R_1, R_2, V_{CC}. (b) Bias circuit redrawn without common ground. (c) Thevenin equivalent circuit for voltage divider bias.

The final Thevenin equivalent circuit (looking back from B – G) consists of R_B' and V_{BB}, where

Example 12-3. Determine the values of R_1 and R_2 in Fig. 12-10 to provide a base bias current I_B = 0.3 mA, so as to locate the operating point at $I_{CQ} = 18$ mA, $V_{CEQ} = 4.25$ V, the same as the conditions in Fig. 12-8(b) in Example 12-2. Make the stability factor S=10 ($\beta = 100$).

Solution. In the base circuit, using the equivalent bias circuit, Fig. 12-11(c):

$$V_{BB} = I_B R_B' + V_{BE} + I_C R_e$$

$$V_{BB} = 0.3 \times 10^{-3} \times R_B' + 0.2 + 18 \times 10^{-3} \times 0.1 \times 10^3$$

$$= 0.3 \times 10^{-3} \times R_B' + 2.$$

This is as far as one can go without choosing a value for R_B'. Since R_e has been chosen as 0.1 K ohm, the selection of R_B' will determine the stability factor S. When R_B' has been selected, the required bias resistors can be determined.

From the curve of S against $\frac{R_B}{R_e}$ for $\beta = 100$ (Fig. 12-9), a reduction of S from 101 to 10 would be accomplished by making $\frac{R_B}{R_e} \approx 10$.

$$\therefore R_B = 10\, R_e = 10 \times 0.1 \text{ K} = 1 \text{ K} = R_B'$$

$$\therefore V_{BB} = 0.3 \times 10^{-3} \times 1 \times 10^3 + 2$$

$$= 0.3 + 2 = 2.3 \text{ V}.$$

Now,

$$R_B' = \frac{R_1 R_2}{R_1 + R_2} \text{ and } V_{BB} = V_{CC} \times \frac{R_2}{R_1 + R_2}.$$

$$\text{Thus: } 1 \text{ K} = \frac{R_1 R_2}{R_1 + R_2} \text{ and } 2.3 = 12 \times \frac{R_2}{R_1 + R_2}.$$

$$\therefore (R_1 + R_2)\, 1 \text{ K} = R_1 R_2 \text{ and } (R_1 + R_2) 2.3 = 12 R_2.$$

From the second equation:

$$2.3 R_1 + 2.3 R_2 = 12 R_2$$

$$2.3 R_1 = 9.7 R_2$$

$$R_1 = \frac{9.7}{2.3} R_2 = 4.2 R_2.$$

Substituting in the first equation:

$$(4.2 R_2 + R_2)\, 1 \text{ K} = 4.2 R_2 \times R_2$$

$$5.2 R_2 \times 1 \text{ K} = 4.2 R_2 \times R_2.$$

$$\therefore 5.2 \text{ K} = 4.2 R_2.$$

$\therefore R_2 = \frac{5.2 \text{ K}}{4.2} = 1.24$ K ohms and, $R_1 = 4.2 R_2 = 4.2 \times 1.24 \text{ K} = 5.2$ K ohms.

The actual value of S is:

$$S = \frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B'}} = \frac{100 + 1}{1 + \frac{100 \times 0.1}{0.1 + 1}} = \frac{101}{1 + \frac{10}{1.1}}$$

$$= \frac{101}{1 + 9.1} = \frac{101}{10.1} = 10.$$

This would be a very stable circuit against temperature variations. However, the effect of $R_2 = 1.24$ K ohms and $R_1 = 5.2$ K ohms is to provide an effective loading of 1 K ohm in parallel with the ac input and may load the input source since the input resistance to the transistor is usually about 1 K ohm. The loading would be even greater if more bias stabilization were required (smaller S).

The reason for this state of affairs (high loading with S = 10 instead of 2 or 3) is the small value of R_L. The value of R_B' could be made larger, say 2 K ohms, if R_e were increased to 200 ohms with S remaining at 10, but with $R_L = 330$ ohms this would make for an inefficient circuit, since R_e would be almost equal to R_L, and there would be an almost equal amount of dc power lost in each. In a voltage amplifier with $R_L = 5$ K ohms, or 10 K ohms, an S of 3 could easily be achieved without loading the input through using higher values for R_1 and R_2. However, where bias stabilization is important, such as low load resistance values in power amplifiers it can always be accomplished by using the voltage divider circuit with R_1 and R_2. The extra stabilization could not be accomplished to the same degree by using R_e and only one resistor R_B, as used in the circuit of Fig. 12-8(a).

12-9. EMITTER BYPASS CAPACITOR

The stabilization due to the addition of R_e has come about due to the action of dc negative feedback. In order to avoid degenerative ac feedback, R_e must be bypassed by a large capacitor C_e, as shown in Fig. 12-12(c). This is much larger than in a tube circuit; often 50 to 100 μF is necessary. This will be covered in detail in the discussion of frequency response characteristics in Chapter 15.

12-10. SUMMARY OF STABILIZATION CIRCUITS

Fig. 12-12 is a summary of the stabilized circuits, all of which have a bias $I_B = 0.3$ mA. The S factors are shown for comparison, as are the loading effects of the biasing resistors. Various combinations of the circuits are possible. For example, in the second circuit, more flexibility in obtaining the bias point can result from adding a resistor from base to ground. Also, an emitter resistor could be added to the second diagram (using collector-to-base biasing) to increase the stability further.

12-11. ADDITIONAL STABILITY FACTORS

It was mentioned that there are two other factors that contribute to thermal instability, variation in I_B (V_{BE}) and β with temperature. Stabilities due to these quantities are measured by S′ and S″ respectively:

$$S' = \frac{dI_C}{dV_{BE}},\ I_{CO} \text{ and } \beta \text{ held constant,}$$

$$= \frac{-S}{R_B + R_e} \times \frac{\beta}{\beta + 1}\ \frac{\text{amps}}{\text{volts}}. \qquad (12\text{-}16)$$

$$S'' = \frac{dI_C}{d\beta},\ I_{CO} \text{ and } V_{BE} \text{ held constant,}$$

$$= \frac{I_C\, S}{\beta_1(\beta_2 + 1)} \text{ amps,} \qquad (12\text{-}17)$$

where β_1 is at temperature T_1 and β_2 is at temperature T_2.

It can be seen that the methods covered for reducing S (by making $\frac{R_B}{R_e}$ small) also serve to reduce S″. However, S′ is minimized by using a large value of R_B *or* R_e (as well as making S small). Thus, in all cases, the largest practical value of R_e should be used, but a compromise will be necessary in selecting R_B.

In general, reducing S will improve thermal stability due to all effects.

The total change in collector current as a function of all three stability factors is:

$$\Delta I_C = S\Delta I_{CO} + S'\Delta V_{BE} + S''\Delta\beta. \qquad (12\text{-}18)$$

Due to the smaller dependence on I_{CO} for a silicon transistor, manufacturers specify the average range of operation for a germanium transistor as −65°C to +75°C and for a silicon transistor as −65°C to +175°C.

This does not mean the transistors will automatically operate successfully in any circuit over this range. The circuits must be stabilized for this operation, but a germanium transistor would not be used at ambient temperatures exceeding +75°C even if adequately stabilized.

In a silicon transistor, the increase in collector current over a given temperature range is due primarily to V_{BE} (S′) and not I_{CO} (S).

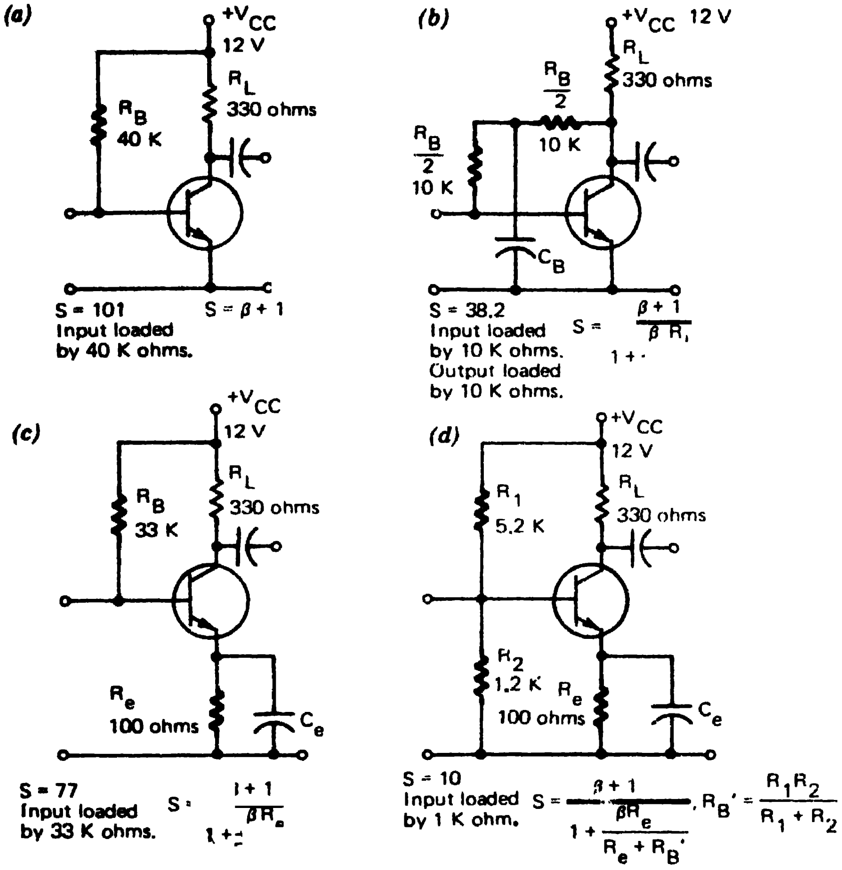

(a)
+VCC
12 V
RB
40 K
RL
330 ohms
S = 101
S = β + 1
Input loaded
by 40 K ohms.
(b)
+VCC 12 V
RL
330 ohms
RB/2
10 K
RB/2
10 K
CB
S = 38.2
Input loaded
by 10 K ohms.
Output loaded
by 10 K ohms.
$S = \frac{\beta + 1}{1 + \frac{\beta R_L}{R_B}}$
(c)
+VCC
12 V
RB
33 K
RL
330 ohms
Re
100 ohms
Ce
S = 77
Input loaded
by 33 K ohms.
$S = \frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B}}$
(d)
+VCC
12 V
R1
5.2 K
RL
330 ohms
R2
1.2 K
Re
100 ohms
Ce
S = 10
Input loaded
by 1 K ohm.
$S = \frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B'}}$, $R_B' = \frac{R_1 R_2}{R_1 + R_2}$

Fig. 12-12. Summary of the stabilized CE amplifier circuits:
(a) Unstabilized—fixed bias.
(b) Collector-to-base bias.
(c) Emitter and fixed bias.
(d) Emitter and voltage divider bias.

However, reducing S will reduce S' and make it stable against variations in V_{BE} due to temperature change. In a germanium transistor the effect of I_{CO} has the dominant influence on the collector current.

12-12. BIAS COMPENSATION

All of the previous circuits used some type of negative feedback to stabilize the operating point. It is also possible to use diodes and thermistors to *compensate* for variations in current, although a mixture of negative feedback and compensation is often used. See Fig. 12-13.

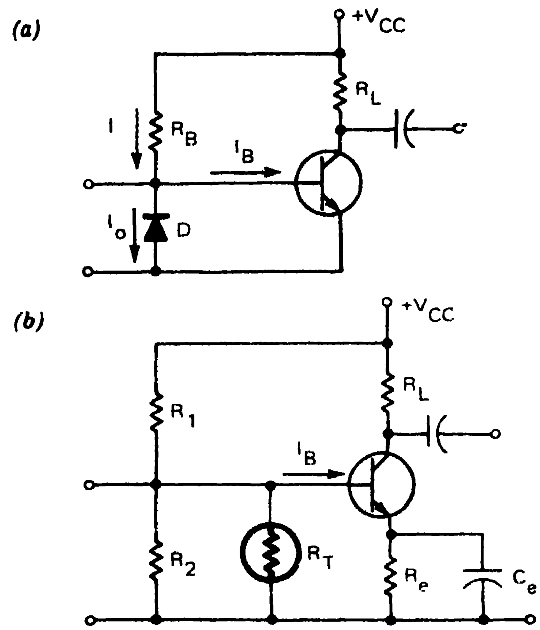


Fig. 12-13. Methods of bias compensation: (a) Use of a reverse-biased diode across the base-emitter junction. (b) Use of a thermistor to reduce forward bias with an increase in temperature. Bias stabilization is also provided by R_e and C_e.

In Fig. 12-13(a), a diode of the same material as the transistor is reverse biased by the base-emitter junction voltage, allowing leakage current I_o to flow through the diode D. If an increase in temperature occurs, causing I_{CO} to increase in the transistor, the leakage current through D will also increase, decreasing the base current I_B (I is approximately constant). This is the required action to keep I_C near constant. Note that this compensation does not need a change in I_C to effect the change in I_B. It is dependent only upon temperature, so theoretically I_o and I_{CO} can track almost equally, and I_C can remain essentially constant.

In Fig. 12-13(b), a thermistor, R_T, which has a negative temperature coefficient is placed in parallel with R_2, and resistance values are selected so that correct bias is applied at normal temperature. An increase in temperature will decrease forward bias, reducing I_B and I_C.

12-13. SUMMARY

(1) The variation in a transistor's dc collector current with temperature in a CE amplifier depends upon β, I_B and I_{CO}, all of which increase with temperature.

(2) The resulting increase in collector current and shift in operating point may cause clipping and distortion of the applied signal, or it may cause thermal runaway.

(3) The stability factor S is defined as the rate of change of collector current I_C with respect to the leakage current I_{CO}, with both β and I_B held constant. This instability factor indicates how much change there is in I_C for a given change in I_{CO} and depends upon the bias circuit used.

(4) The stability factor S for a CE circuit using fixed bias is $\beta + 1$ for either a silicon or germanium transistor.

(5) A common-base circuit has $S = 1$, which is the most stable possible.

(6) Stabilization methods involve the reduction of forward bias with temperature increase, to reduce I_B and hence I_C by some means of negative feedback, in various bias circuits.

(7) Collector-to-base bias employs a resistor R_B connected from collector to base and provides bias stabilization with $S = \dfrac{\beta + 1}{1 + \dfrac{\beta R_L}{R_L + R_B}}$

(8) Collector-to-base bias may not provide much *overall* increase in stability in a silicon transistor circuit since the reduced value of R_B will allow a larger increase in I_B with temperature than with fixed bias.

(9) Emitter bias consists of the addition of a resistor R_e in the emitter lead, bypassed by a large capacitor C_e.

$$\frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B}}.$$

(10) Maximum reduction of S with emitter bias is accomplished by the use of the minimum value of $\frac{R_B}{R_e}$ that is practical.

(11) Voltage divider bias with emitter bias will permit the selection of any operating point with any desired value of S, but usually with increased loading.

$$S = \frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B'}} \qquad R_B' = \frac{R_1 R_2}{R_1 + R_2}.$$

(12) Additional stability factors are S′ and S″, which indicate the dependence of I_C upon I_B and β, respectively.

(13) In a silicon transistor, the effect of I_B(S′) is predominant over I_{CO} (S), but the reduction of S will serve to reduce both S′ and S″.

(14) Bias compensation involves the use of diodes or thermistors to compensate for the variations in current instead of using a negative feedback method.

QUESTIONS

Q12-1. What three factors contribute to thermal instability? Which one has the least effect in silicon transistors? Why?

Q12-2. What do you understand by the term, "thermal instability"? Explain in your own words.

Q12-3. What is the underlying principle of providing thermal stabilization by means of bias circuit modifications? How does this differ from the compensation techniques using a diode or thermistor?

Q12-4. What is the definition of the stability factor, S? Why would it seem more reasonable to call this an *instability* factor? Which circuit has the highest S factor? Which one the lowest?

Q12-5. Describe how a CE amplifier with collector-to-base bias provides thermal stabilization when a temperature rise tries to increase the average collector current.

Q12-6. What are the disadvantages of collector-to-base bias? Can they be overcome?

Q12-7. Sketch a CE amplifier employing a PNP transistor and a resistor in the emitter lead, similar to Fig. 12-7. Describe the stabilizing action of R_e as a temperature increase takes place. What part does the ratio $\frac{R_B}{R_e}$ play in the stabilization?

Q12-8. What is a drawback of emitter bias? What modification should be made if an emitter bias resistor is added to an unstabilized amplifier?

Q12-9. What essentially is gained by the use of a voltage divider bias in the base (Fig. 12-10) compared with a single base bias resistor (Fig. 12-8)?

Q12-10. In selecting the components R_1 and R_2 in Fig. 12-10, with a given R_e, what decision must be made to complete the selection?

Q12-11. If a CE amplifier has a voltage divider base bias circuit, how will the loading on the input compare with a single resistor bias circuit if both have the same value of emitter resistor and base bias current? (See Fig. 12-12.)

Q12-12. What two equations indicate that a reduction in S also serves to increase thermal stability against variations in V_{BE} and β? Does this mean R_B/R_e should be made as small as possible in all cases? Explain.

Q12-13. What is the total change in collector current dependent upon as the temperature changes?

PROBLEMS

P12-1. A 2N404A germanium transistor has an I_{CBO} (I_{CO}) of 5 μA at an ambient temperature of 25°C. Assuming its β (h_{FE}) is 50 and that its base-emitter bias voltage in the CE connection is 0.2 V, use the curves in Fig. 12-1 to determine the collector current at:

(a) a temperature of 25°C.
(b) a temperature of 100°C.

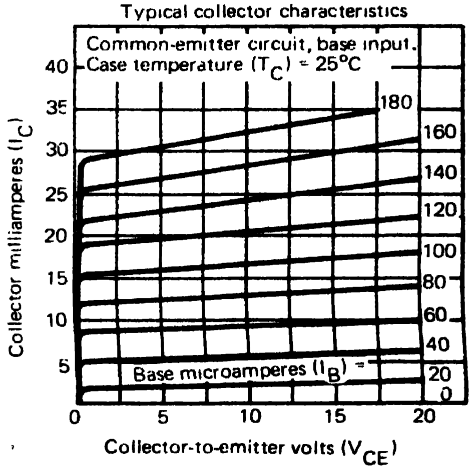
(Use I_{CO} at 95°C.)

P12-2. A silicon transistor is to be used in a CE amplifier circuit like that in Fig. 12-3(a), with V_{CC} = 20 V and R_L = 800 ohms. Using the given collector characteristics, determine:

(a) The dc (and ac) load line.
(b) The value of R_B to provide a base bias current of 80 μA.
(c) The operating point: V_{CE}, I_C.
(d) The value of h_{FE} at the Q point:

$$h_{FE} = \frac{I_C}{I_B}.$$

(e) The stability factor of the circuit, S.
(f) The new value of R_B if collector-to-base bias is to be used to provide the same operating point as in (b). (See Fig. 12-4(a).)
(g) The new stability factor of the circuit with collector-to-base bias.
(h) Whether or not this type of bias makes I_C relatively insensitive to changes in β. Is it any better than fixed bias? By how much?
(i) The values of resistors and bypass capacitor, C_B, in Fig. 12-5 if the lowest frequency to be handled is 10 Hz.



Characteristics of the silicon transistor for P12-2 through 5.

P12-3. If the same transistor as in P12-2 is to be used with emitter bias as in Fig. 12-6 with V_{CC} = 20 V and R_L = 800 ohms, what would have to be the value of R_e to give approximately the same base and collector currents as in P12-2(c) and to provide:

(a) An S factor of 100? (Use $\beta = h_{FE}$ from P12-2(d).)
(b) The same S factor as with collector-to-base bias, in P12-2(g)?

What do you conclude regarding an emitter resistor to provide good thermal stability for this transistor, load, and operating point ($I_B = 80\ \mu A$)?

P12-4. If $R_e = 200$ ohms is used in the amplifier of Fig. 12-8 with $R_L = 800$ ohms, $V_{CC} = 20$ V and the above transistor, determine:

(a) The dc load line.
(b) The value of R_B to provide a base bias current of 80 μA.
(c) The value of the stability factor, S. Does this substantiate the results of P12-3?

P12-5. The amplifier of P12-4 is to be modified so that a voltage divider circuit will be used to bias the base as in Fig. 12-10, instead of using a single resistor, R_B. If $R_L = 800$ ohms, $V_{CC} = 20$ V, $R_e = 200$ ohms, calculate:

(a) The value of R_B' to provide an S factor of 20. Use Eq. 12-15 and the value of β (h_{FE}) from P12-2(d).
(b) The values of R_1 and R_2 to provide a base bias of 80 μA.
(c) The additional loading on the input signal due to R_1 and R_2.
(d) The value of bypass capacitor, C_e, to provide $X_{Ce} \leq 0.1\ R_e$ at the lowest frequency, 100 Hz in this case.

P12-6. Repeat P12-2 using a germanium SK3014 power transistor with $V_{CC} = 30$ V, $R_L = 6$ ohms, base bias $I_B = 10$ mA.

P12-7. Repeat P12-4 using a germanium SK3014 power transistor with $V_{CC} = 30$ V, $R_L = 6$ ohms, $R_e = 1$ ohm, $I_B = 10$ mA.

P12-8. Repeat P12-5 using the SK3014 transistor with $V_{CC} = 30$ V, $R_L = 6$ ohms, $R_e = 1$ ohm, $I_B = 10$ mA. (Omit part (d); R_e is usually left unbypassed in power amplifiers to reduce distortion.)

P12-9. Calculate the percent second harmonic distortion and the output power in R_L assuming a 20 mA p-p signal current into the base for:

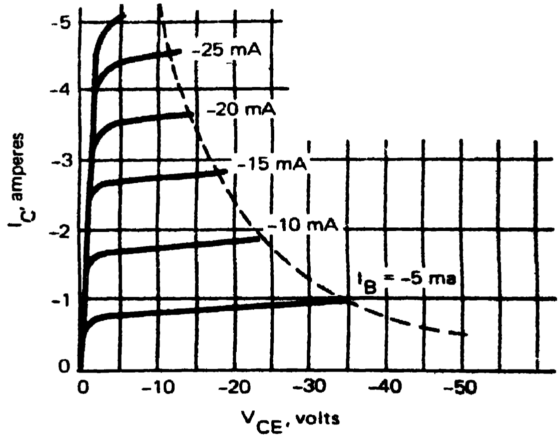
(a) Problem P12-6.
(b) Problem P12-8.

P12-10. A silicon transistor is used in the circuit of Fig. 12-10 with $R_L = 800$ ohms, $V_{CC} = 20$ V, $R_e = 200$ ohms, $R_B = R_1 \| R_2 = 4.4$ K ohms. The characteristics of the transistor at 25°C and 100°C are given in the following table:

25°C: $I_{CO} = 10$ nA, $\beta = 100$, $V_{BE} = 0.6$ V.

100°C: $I_{CO} = 2\ \mu$A, $\beta = 160$, $V_{BE} = 0.41$ V.

(a) Calculate the stability factor S at 25°C.
(b) Calculate S′ at 25°C using Eq. 12-16.
(c) Calculate S″ at 100°C using Eq. 12-17 if the quiescent collector current is 10 mA.
(d) Determine the total change in collector current when the temperature increases from 25°C to 100°C, using Eq. 12-18. (Note that both S′ and ΔV_{BE} are negative.) What is the new value of collector current?



Characteristics of the SK3014 power transistor.

(e) Assume now that a germanium transistor is substituted for the silicon transistor. Its characteristics at 25°C and 75°C are given in the following table:

25°C: $I_{CO} = 10\,\mu A$, $\beta = 100$, $V_{BE} = 0.25$ V.

75°C: $I_{CO} = 0.24$ mA, $\beta = 160$, $V_{BE} = 0.06$ V.

Determine the change in collector current as the temperature increases from 25°C to 75°C if the collector current at 25°C is also 10 mA. What can you conclude regarding the upper temperature of operation of a silicon transistor compared with a germanium transistor for the same change in collector current?

(f) What would the change in collector current have been in each case if neither circuit has its bias point thermally stabilized? That is, if fixed bias were used with R_B = 200 K ohms.

13

THERMAL STABILITY II: TRANSISTOR DISSIPATION

13-0. INTRODUCTION

This chapter examines more closely the problem of thermal runaway, in which internal heating in a transistor (which is due to the difference between the power supplied to the transistor and the power delivered to the load) can cause an increase in collector current. This in turn can increase the collector junction temperature and produce a "snowball" effect called thermal runaway, in which the collector junction may become permanently damaged. The power dissipation of a transistor depends upon its physical size and the way in which it is mounted, so that a discussion of heat sinks and their thermal characteristics is also included. Finally, the cycle of heating and cooling that a transistor experiences as equipment is turned on or off is considered under the topic of thermal fatigue and requires the definition of a transistor's thermal-cycling rating.

13-1. TRANSISTOR DISSIPATION

The maximum power that a transistor may be capable of dissipating may vary from a few milliwatts to several hundred watts, depending upon its construction and heat sinks. (An experimental transistor manufactured by RCA can dissipate 800 W, but 50 to 100 W is typical for a power transistor.) The power handling capability for a given transistor is dependent upon two things: (1) the maximum temperature at which the junction of the device is permitted to operate and (2) the actual ambient temperature expected in operation.

For silicon transistors, the maximum permissible temperature that the collector-to-base junction can be subjected to is in the range of 150°C to 225°C; for germanium transistors, the range is 60°C to 100°C.

For example, a 2N217 germanium transistor has a junction operating temperature range of -65°C to 71°C, although its temperature range in storage has an upper limit of 85°C. (When being soldered into a circuit, its lead temperature may be allowed to reach 255°C for not more than 10 seconds.) The transistor dissipation is listed as 165 mW at ambient temperatures (same as the case temperature for this transistor) up to 25°C (77°F). However, if the ambient temperature exceeds 25°C in a particular environment, then a power-temperature derating curve is used to determine the maximum dissipation capability at elevated temperatures. A typical power-temperature derating chart is shown in Fig. 13-1, which applies to the majority of transistors. It consists of graphs of percent permissible dissipation against transistor case temperature.

Note: In a power transistor, the case temperature may be higher than the ambient temperature. The case or mounting flange temperature must be used when derating the transistor, not the ambient. However, in signal-type transistors (having a maximum dissipation of only a few hundred milliwatts), the case and ambient temperatures are often considered to be the same.

Let us illustrate the use of Fig. 13-1 with a numerical example.

Example 13-1. Determine the dissipation capability of a 2N217 germanium signal transistor at an ambient (case) temperature of 50°C (122°F) if its maximum dissipation is 165 mW at an ambient temperature of 25°C. Given: Maximum permissible operating

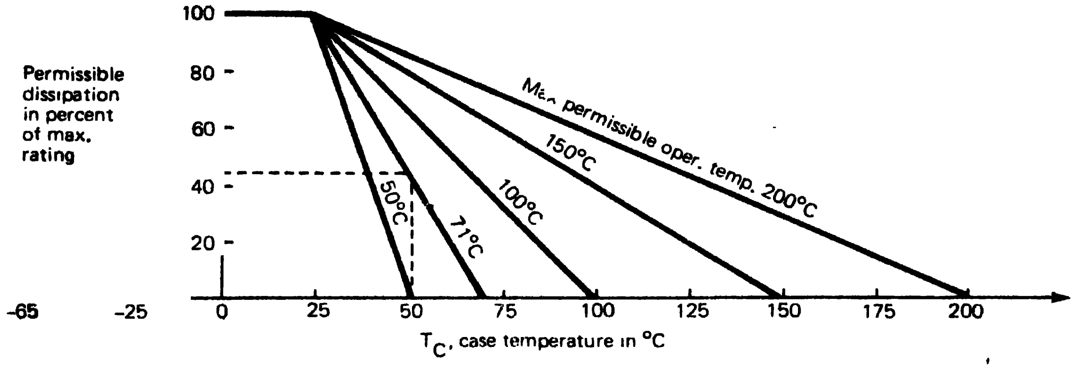


Fig. 13-1. Typical power-temperature derating curves showing maximum permissible percentage of the maximum rated dissipation as a function of case temperature for different transistor maximum operating temperatures.

temperature of the 2N217 is 71° C.

Solution. From the 50° C point on the horizontal ambient (case) temperature axis, draw a vertical line upward to intersect the 71° C line (since the transistor's maximum permissible operating temperature is 71° C). Draw a line horizontally from this point of intersection to the vertical axis and read off 45% of maximum rating. Then the dissipation capability at 50° C ambient is $\frac{45}{100}$ of 165 mW = 74 mW.

Thus, from Fig. 13-1, a 2N217 has zero dissipation capability at 71° C. However, not all transistors conform to this derating curve. For example, a 2N104, which has a maximum permissible operating temperature of 70° C and and a dissipation of 150 mW at 25° C, is capable of safely dissipating 30 mW at 70° C, instead of zero.

13-2 SIGNIFICANCE OF THE QUIESCENT POINT IN THERMAL RUNAWAY

It is possible to determine from an examination of the operating point and ac load line whether a transistor amplifier is inherently stable against thermal runaway. Consider the load line and operating points shown in Fig. 13-2.

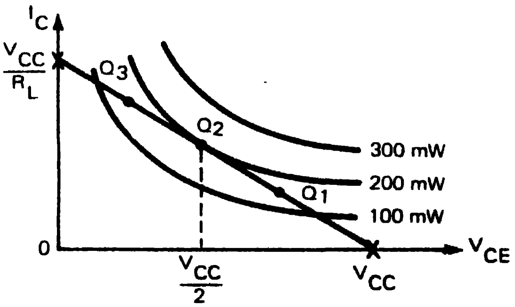


Fig. 13-2. Curves showing the significance of quiescent operating points. With bias at Q_3 the circuit is inherently stable against thermal runaway. With bias point at Q_1, an increase in temperature will increase transistor dissipation and lead to possible thermal runaway.

At Q_2, the ac load line is tangent to the 200 mW dissipation curve, and the dissipation will be 200 mW at quiescent conditions. If the bias point is at Q_3 and an increase in collector current takes place for some reason (such as internal heating or ambient temperature increase), then the dissipation (which was somewhere between 100 mW and 200 mW) will decrease. That is, the operating point will move along the load line toward the 100 mW hyperbola and *away* from the 200 mW curve.

At Q_1 (which has the same dissipation as at Q_3), however, an increase in collector current will move the operating point *toward* the 200 mW dissipation curve and will increase dissipation in the transistor. Although this may not be automatically harmful and lead to thermal runaway, the circuit is not *inherently* stable and must be checked for suitable stabilization. Thus, if the collector-emitter voltage for the Q point is less than $\frac{V_{CC}}{2}$ (and is within the maximum dissipation capability of the transistor), then the Q point is in a safe region where an increase in collector current will decrease transistor dissipation and prevent thermal runaway.

One case very susceptible to thermal runaway is the transformer-coupled amplifier which will be studied in detail in Chapter 19. It has only a small dc resistance in the collector circuit (that of the transformer primary winding and emitter resistor if any), as shown in Fig. 13-3.

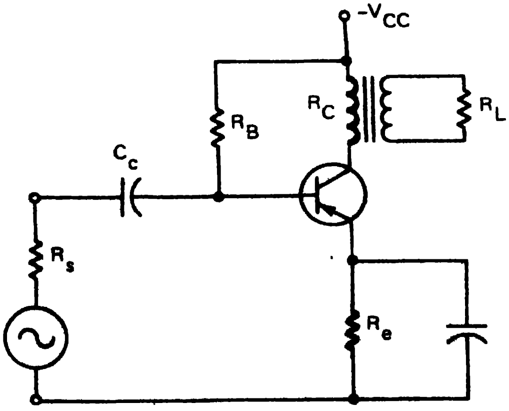


Fig. 13-3. Circuit of a transformer-coupled amplifier in which the total dc resistance in the collector circuit is the sum of R_C and R_e

Here, the Q point may easily be to the right of $\frac{V_{CC}}{2}$, since the dc load line would be drawn almost vertically upward from $V_{CE} = V_{CC}$. In such cases, it is necessary to check the thermal characteristics of the transistor to ensure thermal stability. This is done by examining the *thermal resistance* of the transistor.

13-3. THERMAL RESISTANCE

Consider a transistor being used in a circuit where the ambient temperature of the air around the transistor is represented by T_A degrees centigrade, and the temperature of the collector-base junction of the transistor is T_J. Due to heating within the transistor, T_J is higher than T_A. See Fig. 13-4.

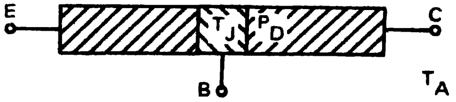


Fig. 13-4. Schematic diagram of a transistor, showing the junction temperature T_J, the ambient temperature T_A, and the power dissipation P_D.

Now it is obvious that the greater the power dissipated in the transistor, P_D, the greater will be the temperature difference between the junction and the ambient temperature. That is, the temperature rise at the collector junction above ambient is proportional to the power dissipated in the transistor.

Stated mathematically, this means that

$$T_J - T_A \propto P_D. \tag{13-1}$$

Recalling that a proportionality can be replaced by an equality if a constant of proportionality is introduced, the following results:

$$T_J - T_A = \theta P_D, \tag{13-2}$$

where θ is the constant of proportionality and is called the thermal resistance. Its units of measurement may be found by rearranging for θ:

$$\theta = \frac{T_J - T_A}{P_D} = \frac{^\circ C}{\text{watts}} . \tag{13-3}$$

Thus θ is measured in °C/watt and may be as small as 0.2 °C/watt for a high-power transistor that has an efficient heat sink or up to 1000 °C/watt for small-signal low-power transistors which have no cooling provision.

What does θ mean physically, and why is it good to have a low value of thermal resistance?

Rearranging the above equation for T_J yields:

$$T_J = T_A + \theta P_D . \tag{13-4}$$

This states that the collector junction temperature of the transistor T_J will be higher than the ambient temperature T_A by an amount equal to the product of θ, the thermal resistance, and P_D, the power dissipated in the transistor.

Example 13-2. Given a transistor with $\theta = 10$ °C/watt, working in an ambient temperature of 25°C. What will be the junction temperature if the transistor dissipates (absorbs) 2 watts of power?

Solution

$$\begin{aligned} T_J = T_A + \theta P_D &= 25^\circ C + (10^\circ C/W) \times 2\ W \\ &= 25^\circ C + 20^\circ C = 45^\circ C. \end{aligned}$$

In other words, a θ of 10°C/watt means that the junction temperature will increase 10°C for every watt of power dissipated (converted into heat) in the transistor. Evidently, a lower θ, say 1°C/watt, means a lower temperature rise (1°C) per watt of power dissipated. Put another way, a transistor that has a *lower* thermal resistance can dissipate more power before its junction temperature approaches a damaging level. Thus, θ represents the resistance or opposition to heat flow away from the junction, so that as low a value as possible is desirable.

As an example of a low power transistor, the 2N104, which has a transistor dissipation capability P_D of 150 mW, has its thermal resistance described as 0.4°C/mW, which is equivalent to 400°C/watt, a typical figure for signal transistors.

Junction-to-Case Thermal Resistance

It has been assumed to this point that θ represents the total thermal resistance from the transistor junction to the ambient air, which should properly be referred to as $\theta_{J\text{-}A}$. However, this is not always the value given in transistor manuals, especially for power transistors. For these the thermal resistance given is from junction-to-case, $\theta_{J\text{-}C}$. Thus, we must be specific in referring to the thermal resistance of a transistor. For example, a 2N1701 silicon diffused junction transistor with a P_D of 25 watts @ 25°C case temperature has $\theta_{J\text{-}C} =$ 7°C/W and $\theta_{J\text{-}A} = 100$°C/W. The junction temperature at rated dissipation may be determined from either of the following equations:

$$T_J = T_A + \theta_{J\text{-}A} P_D \tag{13-5}$$

$$T_J = T_C + \theta_{J\text{-}C} P_D . \tag{13-6}$$

Note that the ambient temperature is used with $\theta_{J\text{-}A}$, but the case temperature is used with $\theta_{J\text{-}C}$. These equations imply (for the given numerical values of $\theta_{J\text{-}C}$ and $\theta_{J\text{-}A}$) that, for a given junction temperature and dissipation in the transistor, the temperature of the transistor case will be higher than the ambient temperature. It is for this reason that heat sinks are used on transistors and will be discussed in Section 13-5.

13-4. CALCULATION OF TRANSISTOR DISSIPATION

Quantitatively, the amount of power that may be safely dissipated in a transistor is given by:

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} \qquad (13\text{-}7)$$

or

$$P_D = \frac{T_J - T_C}{\theta_{J\text{-}C}}. \qquad (13\text{-}8)$$

The maximum permissible operating junction temperature T_J is specified in the transistor manual, and T_A will depend upon the specific environment anticipated. Essentially, the second equation is represented by the power-temperature derating curves of Fig. 13-1 as far as determining P_D at a given case temperature, T_C, is concerned.

The more usual problem is one in which the expected ambient temperature T_A is known, as is T_J. But, if $\theta_{J\text{-}A}$ is not given, how can we use the first equation, that is,

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}}?$$

If the thermal resistance from junction to ambient is considered, it is seen to consist of two parts:

$$\theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}A}. \qquad (13\text{-}9)$$

That is, heat dissipated in the junction must make its way to the surrounding air through two series paths; from junction-to-case and then case-to-air.

Taking the 2N1701 as an example, with $\theta_{J\text{-}A} = 100°\text{C/W}$ and $\theta_{J\text{-}C} = 7°\text{C/W}$, then,

$$\theta_{C\text{-}A} = \theta_{J\text{-}A} - \theta_{J\text{-}C}$$

$$(100 - 7)°\text{C/W} = 93°\text{C/watt}. \qquad (13\text{-}10)$$

Obviously, the greatest obstruction to heat flow from junction to air is represented by the large value of $\theta_{C\text{-}A}$, case-to-ambient air.

Example 13-3. Determine the power that a 2N1701 can safely dissipate in free air when $T_A = 50°\text{C}$. Given that $T_J = 200°\text{C}$ and $\theta_{J\text{-}A} = 100°\text{C/W}$.

Solution

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{200 - 50}{100} = \frac{150}{100} = 1.5 \text{ watts}.$$

If the transistor were in an enclosed space, T_A might build up to a value higher than 50°C. This would reduce the differential $T_J - T_A$ and would cause overheating of the junction due to reduced heat transfer from the transistor.

13-5. HEAT SINK THEORY

The question now arises, given a transistor, how may we increase the amount of power, P_D, that it may safely handle? This may be seen by rewriting the expression for P_D and substituting the values for the 2N1701 transistor as follows:

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{T_J - T_A}{\theta_{J\text{-}C} + \theta_{C\text{-}A}} = \frac{T_J - T_A}{7 + 93}. \qquad (13\text{-}11)$$

$\theta_{J\text{-}C}$ is determined by the manufacture of the transistor and how it is located in the case, but $\theta_{C\text{-}A}$ is determined by the surface area of the case or flange and its contact with the air. If the effective surface area of the transistor case could be increased, the resistance to heat flow, or $\theta_{C\text{-}A}$, would be decreased. This can be achieved by the use of a heat sink, of which many versions exist, depending upon the shape and size of the transistor. A heat sink is a relatively large, finned, usually black metallic heat conducting device placed in close contact with the transistor case or flange. (See Section 13-6.) The larger the heat sink, the smaller will be its thermal resistance, $\theta_{HS\text{-}A}$.

Now this thermal resistance $\theta_{HS\text{-}A}$ is not added to $\theta_{C\text{-}A}$ in series, but is instead in parallel with it, and if $\theta_{HS\text{-}A}$ is much less than $\theta_{C\text{-}A}$, then $\theta_{J\text{-}A}$ will be reduced significantly. The thermal resistances involved, representing the paths taken by the heat, can be shown by a resistance diagram, as in Fig. 13-5.

If the heat sink makes good contact with the transistor case, the thermal resistance from case to heat sink, $\theta_{C\text{-}HS}$, can be neglected, leaving $\theta_{HS\text{-}A}$ in parallel with $\theta_{C\text{-}A}$, as illustrated in Fig. 13-5(b).

$$\text{Thus } \theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}A} \| \theta_{HS\text{-}A}. \qquad (13\text{-}12)$$

Let us illustrate how a heat sink can improve the dissipation capability of a transistor with a numerical example.

Example 13-4. Determine the power dissipation capability of a 2N1701 transistor which has been fitted with a heat sink having a thermal resistance $\theta_{HS\text{-}A} = 10°\text{C/W}$. Given: $T_A = 50°\text{C}$, $T_J = 200°\text{C}$, $\theta_{J\text{-}C} = 7°\text{C/W}$, $\theta_{C\text{-}A} = 93°\text{C/W}$.

Solution

$$\theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}A} \| \theta_{HS\text{-}A} = 7 + 93 \| 10$$

$$= 7 + \frac{93 \times 10}{93 + 10} = 7 + 9 = 16°\text{C/W}$$

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{200 - 50}{16} = \frac{150}{16} = 9.4\text{ W}$$

This compares with 1.5 W without a heat sink and the same ambient temperature. Note that the effect of the heat sink is to make $\theta_{J\text{-}A}$ smaller and closer to $\theta_{J\text{-}C}$, with $\theta_{J\text{-}C}$ being achievable only by using an infinite heat sink, that is, $\theta_{HS\text{-}A} = 0$. Thus, $\theta_{J\text{-}C}$ is the limiting factor in the power dissipating capability of a transistor and is the reason that this is the one value that is always given for power transistors. It should be noted that forced-air cooling often accompanies the use of substantial heat sinks since a high value of P_D is still dependent upon as low a value of T_A as possible.

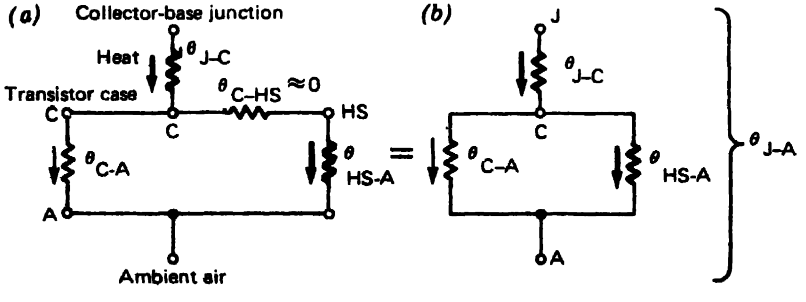


Fig. 13-5. (a) Diagram showing the paths traveled by the heat from junction to ambient, in the presence of a heat sink. (b) Equivalent thermal resistance diagram assuming $\theta_{C\text{-}HS}$ is zero. Note how most of the heat passes through $\theta_{HS\text{-}A}$.

13-6. HEAT SINK HARDWARE AND THERMAL CHARACTERISTICS

Low Power Transistor Types

In some low-power transistors and transistors of the small-signal type, sufficient increased dissipation capability may result from mounting the transistor case directly on the metal chassis. However, this must be done with care, since very often the collector of the transistor is connected to the transistor case to increase heat-dissipation capabilities. Thus, some provision for insulating the case from the chassis (which is usually at ground potential) must be provided, unless a common collector circuit is being employed.

One method is to use a beryllium oxide insulating washer as shown in Fig. 13-6(a). This washer has a good thermal conductivity, and heat transfer from the transistor case to the chassis is improved by the use of a zinc-oxide-filled silicone compound between the washer and the chassis. An insulated clamp over the top of the transistor may be used to increase pressure and help improve thermal dissipation.

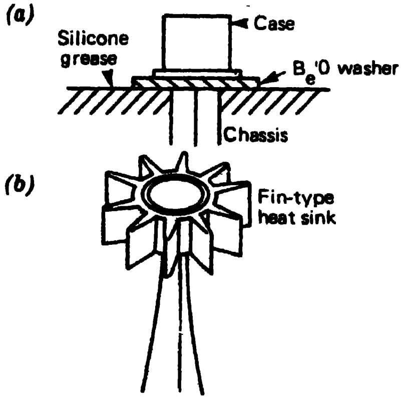


Fig. 13-6. Methods of mounting transistors that are encased in a metal TO-5 package. (a) Mounting the transistor case close to the chassis using a beryllium oxide insulating washer. (b) Using a separate heat sink pressed onto the transistor.

Where the transistor is mounted in teflon sockets that provide no thermal conduction from transistor case to chassis a press-on fin-type of heat sink may be used as shown in Fig. 13-6(b). Typical thermal resistance, $\theta_{HS\text{-}A}$, of such a black anodized heat sink, is in the order of 50° C/W, when designed for a TO-5 transistor package.

Power Transistor Heat Sinks

Popular mounting packages for power transistors are the diamond shaped TO-3 and TO-66 types, which may have dissipations in the order of 100 W. A typical heat sink that can accommodate one TO-3 power transistor and/or two stud mounted diodes is shown in Fig. 13-7. It provides cooling by conduction, convection, and radiation. Although measuring only $4\frac{1}{2}$ in. by $3\frac{1}{16}$ in., it has a thermal dissipation equal to that of a flat aluminum sheet 10 in. × 8 in. × $\frac{1}{8}$ in.

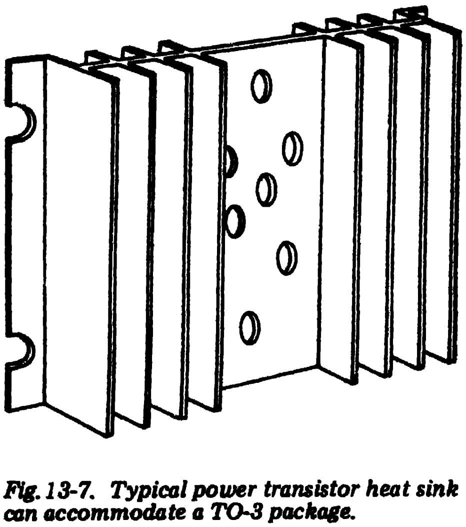


Fig. 13-7. Typical power transistor heat sink that can accommodate a TO-3 package.

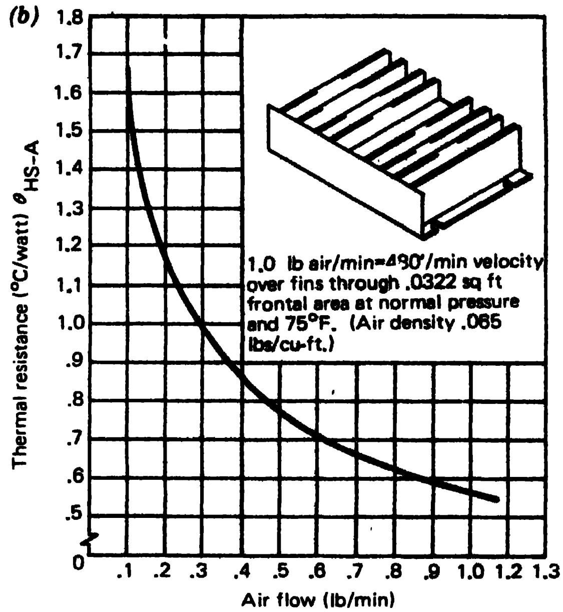
The thermal characteristics for the heat sink in Fig. 13-7 are shown in Fig. 13-8, along with its physical specifications.

(a)

Specifications

Material	Aluminum alloy
Finish	Black
Total surface area	65 sq. in. (approx.)
Thermal resistance	3°C/watt

(b)



(c)

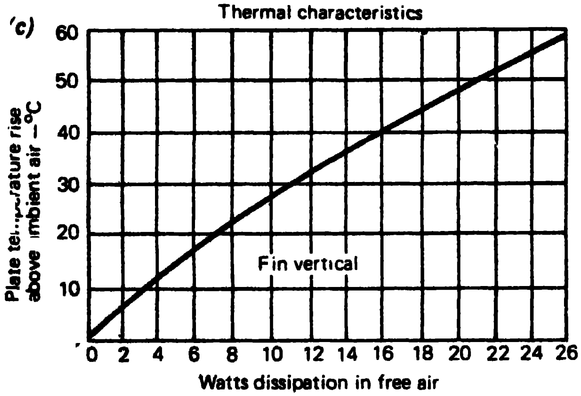


Fig. 13-8. (a) Heat sink specifications showing thermal resistance $\theta_{HS\text{-}A}$ = 3°C/W in free air.
(b) Graph showing the variation of $\theta_{HS\text{-}A}$ with air flow in lb/min.
(c) Graph showing the rise in heat sink temperature above ambient air as a function of watts dissipated in free air. (Courtesy of Motorola, Inc.)

The thermal resistance of this heat sink is 3° C/W in free air, but decreases to approximately 0.6° C/W when exposed to an air flow of 1 pound per minute. See Fig. 13-8(b). This corresponds to a 480 ft/min air velocity over the fins through an area of 0.0322 square feet. (Temperature of air, 75° F or 25° C.)

Fig. 13-8(c) shows the increase in heat sink temperature above ambient air as a function of the power dissipated in the transistor (and heat sink) in free air. This curve will be useful in the discussion of Example 13.5. Let us illustrate the use of the information in Fig. 13-8 with a numerical example.

Example 13-5. A 2N1702 silicon NPN power transistor in a TO-3 package, having a dissipation of 75 W at case temperatures up to 25° C, is mounted in a heat sink having the characteristics shown in Fig. 13-8. Given $T_{J,max}$ = 200° C, $\theta_{J\text{-}C}$ = 2.33° C/W, determine the following:

(a) The power dissipation capability of the transistor when operating in free air at an ambient temperature of 25° C.

(b) The temperature of the heat sink when dissipating the power calculated in (a).

(c) The dissipation capability of the transistor when air at 25° C is forced over the heat sink fins at 0.3 lb/min.

(d) The temperature of the transistor case and heat sink when dissipating the power calculated in (c).

Solution

(a) $P_D = \dfrac{T_J - T_A}{\theta_{J\text{-}A}} = \dfrac{T_J - T_A}{\theta_{J\text{-}C} + \theta_{C\text{-}A} \| \theta_{HS\text{-}A}}$

From the specifications in Fig. 13-8(a), $\theta_{HS\text{-}A} = 3°C/W$ in free air. Since $\theta_{C\text{-}A}$ is not given, we will use:

$$\theta_{C\text{-}A} \| \theta_{HS\text{-}A} \approx \theta_{HS\text{-}A}, \qquad (13\text{-}13)$$

which will be true for large heat sinks having small values of $\theta_{HS\text{-}A}$ compared with $\theta_{C\text{-}A}$ and will provide a value on the safe side.

$$\therefore P_D \approx \frac{T_J - T_A}{\theta_{J\text{-}C} + \theta_{HS\text{-}A}} = \frac{200 - 25}{2.33 + 3} = \frac{175}{5.33} = 32.8 \text{ W.}$$

(b) Although Fig. 13-8(c) does not extend up to a dissipation of 32.8 W, extrapolation of the curve reveals a heat sink temperature rise of approximately 70°C above ambient. Thus, the temperature of the heat sink will be 95°C.

(c) From the curve in Fig. 13-8(b), an air flow of 0.3 lb/min makes $\theta_{HS\text{-}A} = 1°C/W$.

$$\therefore P_D = \frac{T_J - T_A}{\theta_{J\text{-}C} + \theta_{HS\text{-}A}} = \frac{200 - 25}{2.33 + 1} = \frac{175}{3.33} = 52.5 \text{ W.}$$

(d) To determine the case temperature, T_C, recall:

$T_C = T_J - \theta_{J\text{-}C} P_D$ (See Section 13-3.)

$\therefore T_C = 200 - 2.33 \times 52.5 = 200 - 122 = 78°C.$

Note that, once the case temperature is known in (d), then the power temperature derating curve of Fig. 13-1 may be used to obtain the power dissipation capability found in (c). This is done as follows:

In Fig. 13-1 draw a vertical line from a case temperature T_C of 78°C to intersect the 200°C line. Extending a line horizontally provides a percentage dissipation of 70% of the maximum. Thus $P_D = 70\%$ of $75 \text{ W} = \frac{70}{100} \times 75 \cdot 52.5$ W. However, direct use of Fig. 13-1 to solve the problem in (c) or (d) is not possible, since T_C is dependent upon P_D, and P_D may not be determined without knowing T_C! The curves of Fig. 13-1 are good only for determining what the dissipation should be limited to in order to keep the case temperature below some required value.

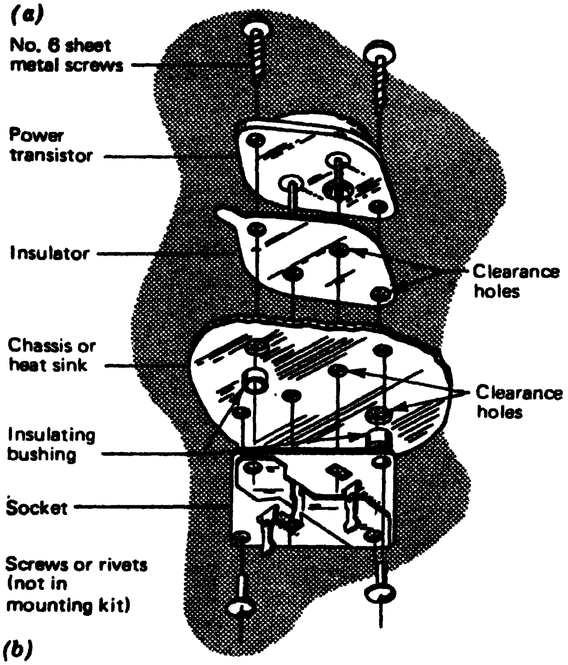
13-7. HEAT SINK MOUNTING

It was mentioned that many power transistors are encased in metal packages of the TO-3 or TO-66 types. These have two leads for the emitter and base, but the case, or the mounting flange of the case, is the collector terminal, so that once again insulation of the case from the heat sink is often necessary. This may be accomplished by the use of an insulating washer as shown in Fig. 13-9(a).

It has been assumed to this point that the thermal resistance from case to heat sink, $\theta_{C\text{-}HS}$, was small enough to be negligible. However, as shown by Example 13-5, the thermal resistance of the heat sink, $\theta_{HS\text{-}A}$, may be less than 1°C/watt, so perhaps the value of $\theta_{C\text{-}HS}$ is significant. Reference to the table in Fig. 13-9(b) confirms this, where $\theta_{C\text{-}HS}$ may be anywhere from 0.4 to 1.45°C/watt, depending upon the washer material. This may be reduced to a range of 0.35 to 0.80 °C/W by the use of silicone grease. Thus, the total thermal resistance from junction to ambient, $\theta_{J\text{-}A}$, must take into account the effect of the mounting hardware and whether silicone grease is used or not:

$$\theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}HS} + \theta_{HS\text{-}A}. \qquad (13\text{-}14)$$

Example 13-6. Calculate the dissipation capability of a 2N1702 transistor operating under forced air cooling conditions on a heat sink



(b)

Insulating washer	Typical thermal resistance (°C/watt)	
	Dry	With DC4*
No insulator	.20	.10
Teflon	1.45	.80
Mica	.80	.40
Anodized aluminum	.40	.35
Mica	.80	.40

*DC4 is Dow Corning No. 4 Silicone Lubricant.

Fig. 13-9. (a) Diagram showing the mounting of a TO-3 or TO-66 transistor package on a heat sink with an insulating washer. (b) Table showing the thermal resistance, $\theta_{C\text{-}HS}$, for various washers, and effect of silicone lubricant. (Courtesy of Motorola, Inc.)

as in part (c) of Example 13-5 for the following conditions:

(a) No insulator, but silicone lubricant used.

(b) Teflon washer, with no lubricant used.

Solution

(a) From Fig. 13-9(b), the thermal resistance $\theta_{C\text{-}HS}$ with no insulator but with lubricant being used is 0.1°C/W.

$$\therefore \theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}HS} + \theta_{HS\text{-}A}$$
$$= 2.33 + 0.1 + 1$$
$$= 3.43°\text{C/W}.$$

$$\therefore P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{200 - 25}{3.43} = \frac{175}{3.43} = 51 \text{ W}.$$

This is substantially the same as in Example 13-5(c).

(b) From Fig. 13-9(b), for Teflon, $\theta_{C\text{-}HS} = 1.45°\text{C/W}$.

$$\therefore \theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}HS} + \theta_{HS\text{-}A}$$
$$= 2.33 + 1.45 + 1$$
$$= 4.78°\text{C/W}.$$

$$\therefore P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{200 - 25}{4.78} = \frac{175}{4.78} = 36.6 \text{ W},$$

which is a substantial reduction from (a).

13-8. SELECTION OF HEAT SINK SIZE

Thus far, we have been concerned with the determination of power dissipation capability for a transistor mounted in a given heat sink. Another type of problem involves the selection of a heat sink to provide a desired dissipation capability in known ambient conditions. This requires the use of the following equation:

$$\theta_{J\text{-}A} = \frac{T_J - T_A}{P_D}. \tag{13-3}$$

Solving for $\theta_{J\text{-}A}$ will permit $\theta_{HS\text{-}A}$ to be determined. The heat sink may be selected from those commercially available which provide a thermal resistance equal to or less than that calculated. Or, some idea of the size of heat sink or area of chassis that is necessary may be obtained from the nomograph in Fig. 13-10. A straight line drawn horizontally from the thermal resistance required (using the column corresponding to the material, thickness, and mounting position in question) to the left hand column will provide the necessary area.

It is interesting to note that a power transistor advertised as having a certain dissipation capability at a case temperature of 25° C will never be able to safely dissipate this power when operated at room temperatures without an *infinite* heat sink. That is, one that has both $\theta_{C\text{-}HS}$ and $\theta_{HS\text{-}A}$ equal to zero, is needed. For example, a 2N1702 transistor has a maximum dissipation of 75 W at T_C = 25° C. Its $\theta_{J\text{-}C}$ = 2.33° C/W.

Thus:

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{T_J - T_A}{\theta_{J\text{-}C} + \theta_{C\text{-}HS} + \theta_{HS\text{-}A}}$$

$$= \frac{200 - 25}{2.33 + 0 + 0} = \frac{175}{2.33}$$

$$= 75 \text{ watts.}$$

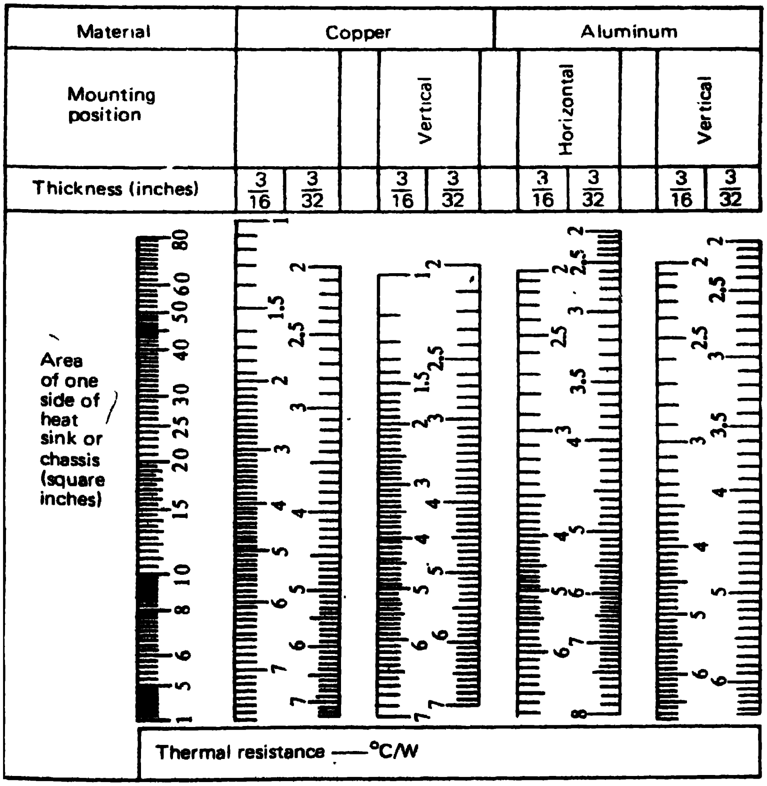


Fig. 13-10. Thermal resistance as a function of heat-sink dimensions. (Courtesty of Electronic Design, August 16, 1961.)

Obviously such a heat sink is not possible, so that the only way a 2N1702 can safely dissipate 75 watts is for the ambient temperature to be lower than 25°C (room temperature). This means that refrigerated air would be necessary to achieve the maximum rated dissipation.

13-9. CONDITIONS FOR THERMAL STABILITY

We have seen from Section 13-2 how an amplifier is inherently stable if the Q point is located so that $V_{CE} < \frac{V_{CC}}{2}$ on the ac load line. That is, an increase in I_C will change the Q point so that transistor dissipation decreases, and thermal runaway is not a danger. However, if this condition is not met (as in many transformer-coupled amplifiers), how can we be sure of thermal stability from other considerations, such as a sufficiently large transistor dissipation capability or heat sink?

It may be shown that if the rate at which heat is released in the transistor is not to exceed the rate at which heat can be dissipated, then:

$$\left[V_{CC} - 2I_{CQ}(R_e + R_C)\right](S)(0.07 I_{CO}) < \frac{1}{\theta_{J\text{-}A}}, * \qquad (13\text{-}15)$$

where

V_{CC} is the collector supply in volts,

I_{CQ} is the quiescent collector current in amps,

R_e is the emitter resistor in ohms,

R_C is the collector circuit resistance in ohms,

S is the stability factor of the circuit,

I_{CO} is the leakage current in amps (I_{CBO}),

$\theta_{J\text{-}A}$ is the thermal resistance of the transistor in °C/W,

0.07 is the 7% increase per °C in the leakage current for either germanium or silicon.

From this equation, the maximum allowable value of $\theta_{J\text{-}A}$ may be determined. (In addition, it is necessary that the transistor have a sufficiently high dissipation capability.) However, it is evident that if S is made low enough this will be the most direct contribution to making the left hand side of Eq. 13-15 less than $\frac{1}{\theta_{J\text{-}A}}$. Consider the following numerical example.

Example 13-7. Let it be required to: (a) check whether the circuit of Fig. 13-11 is inherently thermally stable, (b) specify, if necessary, what is required to make it stable through the use of a heat sink, and (c) determine the necessary dissipation capability of the transistor when operating at a 25°C ambient temperature.

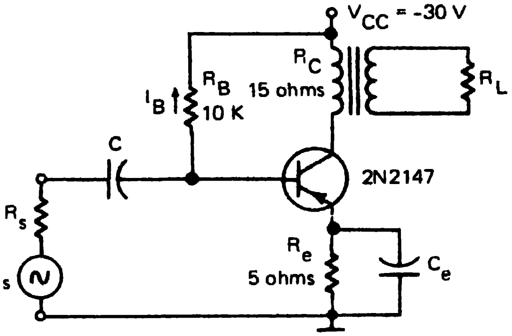


Fig. 13-11. Transformer-coupled power amplifier to be checked for thermal stability.

---

*Millman, J., and Halkias C., *Integrated Electronics,* McGraw-Hill, New York, 1972, Section 9-10.

Solution. From the transistor manual, for a 2N2147 germanium transistor: $\beta = 100$ to 300, $I_{CO} = 1$ mA @ $V_{CB} = 40$ V.

(a) Let us assume $\beta = 100$. We must first find the quiescent collector current, I_{CQ}, in order to determine the quiescent voltage from collector to emitter. This is done by considering the following two equations:

$$I_C = \beta I_B + (\beta + 1)I_{CO} \approx \beta(I_B + I_{CO}) \qquad (13\text{-}16)$$

$$V_{CC} = I_B R_B + V_{BE} + I_C R_e \approx I_B R_B + I_C R_e. \qquad (13\text{-}17)$$

Substituting values for β, I_{CO}, V_{CC}, R_B and R_e yields

$$I_C = 100 I_B + 100 \times 10^{-3}$$

$$30 = 10^4 I_B + 5I_C.$$

Solving these two equations for I_B and I_C gives

$$I_B = 2.8 \text{ mA}, \; I_C = I_{CQ} = 0.38 \text{ A}.$$

The quiescent voltage from collector to emitter is

$$V_{CE} = V_{CC} - I_C(R_e + R_C)$$
$$= 30 - 0.38(5 + 15) = 22.4 \text{ V}.$$

Since this is greater than $\frac{V_{CC}}{2} = \frac{30}{2} = 15$ V, the circuit is not inherently stable, and the transistor's thermal resistance must be checked using Eq. 13-15.

(b) The stability factor S is required:

$$S = \frac{\beta + 1}{1 + \frac{\beta R_e}{R_e + R_B}} = \frac{100 + 1}{1 + \frac{100 \times 5}{5 + 10{,}000}} = 96.2.$$

Let us now determine the value of the left hand side of Eq. 13-1.

$$\big(V_{CC} - 2I_C(R_e + R_C)\big)\,(S)(0.07 I_{CO})$$
$$= (30 - 2 \times 0.38(5 + 15))96.2 \times (0.07 \times 1 \times 10^{-3}) = 0.0996$$

$$\therefore 0.0996 < \frac{1}{\theta_{J\text{-}A}} \text{ or } \theta_{J\text{-}A} < \frac{1}{0.0996}$$
$$< 10.1 \text{ °C/watt.}$$

Thus, the transistor must have a junction-to-ambient thermal resistance, $\theta_{J\text{-}A}$, less than 10.1°C/watt.

Checking the 2N2147, it has a $\theta_{J\text{-}C}$ of 1.5°C/W, but $\theta_{J\text{-}A}$ is not specified. In order to insure that $\theta_{J\text{-}A}$ does not exceed 10.1°C/W, a heat sink should be used having a thermal resistance $\theta_{HS\text{-}A}$ not exceeding 10.1 – 1.5 or 8.6°C/watt, including the mounting hardware. Since the 2N2147 is mounted in a TO-3 package, the heat sink referred to in Fig. 13-8 could be used and would be more than adequate even in free air at 25°C.

(c) The 2N2147 transistor has a maximum junction operating temperature, $T_J = 100$°C Assuming $\theta_{J\text{-}A} = 10$°C/W, the maximum dissipation capability at $T_A = 25$° C is given by:

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} = \frac{100 - 25}{10} = \frac{75}{10} = 7.5 \text{ W}.$$

Compare this with the quiescent dissipation:

$$P = V_{CE} I_{CQ} = 22.4 \times 0.38$$
$$= 8.5 \text{ W}.$$

This is slightly in excess of the permitted value, although the average dissipation over a full cycle when a signal is applied is less than 8.5 W. (This will be discussed in Chapter 19 on power amplifiers.) However, it would be highly desirable to make sure $\theta_{J\text{-}A}$ is less than 10°C/W, for example 8°C/W or less, so that there is some margin of safety should the ambient temperature increase for any reason.

It should be noted, however, that power transistors are not specified for power dissipation capability at ambient temperatures, but, instead, case temperatures, T_C, or mounting flange temperature, T_{MF}. Thus, the question remains, what should be the dissipation capability of the transistor in terms of case temperature?

From Section 13-3, we found

$$T_J = T_C + \theta_{J\text{-}C} P_D \tag{13-6}$$

$$\therefore T_C = T_J - \theta_{J\text{-}C} P_D.$$

Since at quiescence $P_D = 8.5$ W, and $\theta_{J\text{-}C} = 1.5°$C/W, then

$$T_C = 100 - 1.5 \times 8.5$$

$$= 100 - 12.8 = 87°\text{C}.$$

Thus, the transistor must be able to safely dissipate 8.5 W at a case temperature of 87°C. (More specifically, the transistor case temperature must not exceed 87°C for this transistor to be able to safely dissipate 8.5 W.)

Checking the thermal characteristics of the 2N2147 transistor, we find that its dissipation is 12.5 W at mounting flange (case) temperatures up to 81° C and it must be derated linearly 0.66W/°C above this temperature.

Thus, at 87° C its dissipation capability is

$$12.5 - (87 - 81)°\text{C} \times 0.66\ \text{W/°C}$$

$$= 12.5 - 6 \times 0.66 = 12.5 - 4 = 8.5\ \text{W},$$

which just equals the power dissipated in it under quiescent conditions. Note that a heat sink that makes $\theta_{J\text{-}A} = 8.8°$C/W is necessary to provide this dissipation capability.

Thus, for this application, a 12.5 W transistor at $T_C = 81°$C is needed. Alternatively, we could have specified the dissipation capability at a case temperature $T_C = 25°$C by assuming the use of another transistor that obeys the curves in Fig. 13-1. For a maximum operating temperature of 100°C, a case temperature of 87° C corresponds to 20% of the maximum dissipation at 25°C. Thus, the dissipation capability of this transistor at a case temperature of 25° C would have to be 5 × 8.5 W, or 42.5 W!

It should be concluded that transistors must be chosen for their dissipation capability very carefully, and not solely on their quiescent dissipation in a circuit.

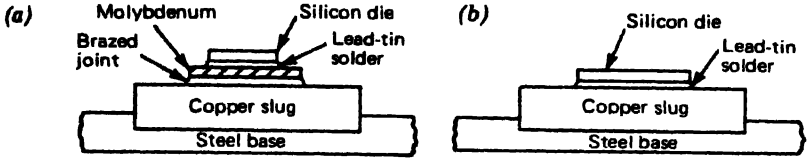
13-10. THERMAL FATIGUE

We have seen from the previous examples how the case temperature of a power transistor may rise 50° to 100° C above ambient, even with a heat sink, when dissipating power near its rated value. This extreme temperature change can cause severe stresses to be set up at the interface between the transistor proper and the header on which it is mounted. Such cyclic stresses produced by the repeated turning on and off of equipment, called thermal cycling, can cause failure of the transistor due to metal fatigue. It is caused basically by the difference in thermal expansion coefficients of the two materials, the silicon of the transistor and the copper header to which it is soldered. See Fig. 13-12(b).

The stresses may be reduced by the use of a molybdenum slab inserted between the two materials, as in Fig. 13-12(a). It has a coefficient of expansion intermediate between silicon and copper and thus provides an expansion matching function. Another method reduces the impurities in the lead which is used for soldering the transistor semiconductor material to the slab. This reduces fatigue failure in the solder due to continuous flexing under thermal cycling.

Thermal-Cycling Rating

The thermal-cycling capability of the transistor — the number of cycles that the transistor



Material	Thermal coefficient of expansion X 10^{-6} per °C
Silicon	4.2
Molybdenum	4.9
Copper	17.0
Steel	11.0

Fig. 13-12. (a) Cross-section of a transistor that uses a molybdenum expansion matcher between pellet and header. (b) Cross-section of a transistor in which pellet is soldered directly to copper. (Courtesy of RCA Corp.)

may be subjected to before failure — is much increased by these methods, often by a factor of at least ten. For example, a group of 2N3773 transistors on a copper header had a 10% failure rate after only 1,000 cycles, but had a thermal-cycling rating of 9,600 cycles when molybdenum was used. In each case, the power dissipated was 85 W and caused a 42° C change in case temperature.

Typical values of thermal-cycling rating required for various applications are shown in Table 13-1 and are determined primarily by the minimum equipment life required and nature of operation.

RCA has started including in technical data thermal-cycling ratings for all silicon power transistors announced since Jan. 1, 1971. However, a chart that may be used to predict the thermal-cycling capability of a power transistor, with $\theta_{J\text{-}C} = 1.5^\circ$ C/W, is shown in Fig. 13-13 and is a function of power dissipation and case temperature change.

Normally, the chart is used to determine the maximum permissible case temperature change from a knowledge of the total power dissipation and the thermal-cycling rating required for a given application. Then a heat sink of the correct dimensions may be selected to keep the change in case temperature within this maximum value and thus avoid transistor failure due to thermal fatigue.

13-11. SUMMARY

(1) The dissipation capability of a transistor at elevated ambient temperatures is reduced by an amount dependent upon the case temperature and the maximum permissible junction operating temperature. This is usually determined by means of a power-temperature derating chart.

(2) If the quiescent point in a CE amplifier is chosen so that $V_{CE} < \frac{V_{CC}}{2}$, the amplifier is inherently stable against thermal runaway. That is, an increase in temperature will cause a reduction in transistor dissipation.

(3) The thermal resistance of a transistor, θ, is a measure of the opposition to heat flow away from the transistor junction, and is measured in °C/watt.

(4) The power that a transistor may safely dissipate may be determined from

TABLE 13-1 THERMAL CYCLING REQUIREMENTS FOR TYPICAL APPLICATIONS OF POWER TRANSISTORS (COURTESY OF RCA CORP.)

APPLICATION	CIRCUIT	P_T (W)	ΔT_C (°C)	MINIMUM EQUIPMENT LIFE REQUIRED (YEARS)	TYPICAL THERMAL-CYCLING RATING REQUIRED (CYCLES)
Auto radio audio output	Class A Class AB	8 2	75 45	5 5	5,000 5,000
Power supply	Series regulator Switching regulator	50 15	65 65	5 5	5,000 5,000
Hi-fi audio amplifier	Class AB	35	50	5	5,000
Computer power supply	Series regulator	50	65	10	10,000
Computer peripheral equipment	Solenoid driver	5	5	10	1.3×10^8
Television	Vertical output Audio output	10 8	75 75	5 5	5,000 5,000
Sonar modulator	Linear amplifier	100	55	10	144×10^3

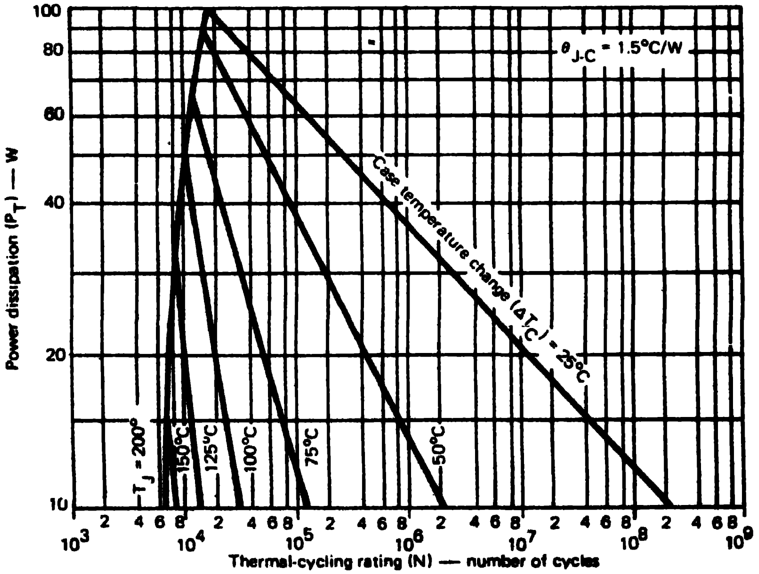


Fig. 13-13. Typical thermal-cycling rating chart as a function of power dissipation and case temperature change. (Courtesy of RCA Corp.)

$$P_D = \frac{T_J - T_A}{\theta_{J\text{-}A}} \quad \text{or} \quad P_D = \frac{T_J - T_C}{\theta_{J\text{-}C}},$$

where $\theta_{J\text{-}C}$ is the junction-to-case thermal resistance, and T_C is the temperature of the case of the transistor.

(5) A heat sink is a metallic heat-conducting device placed in close contact with a transistor to increase the dissipation capability of the transistor by reducing the total thermal resistance from junction to ambient, $\theta_{J\text{-}A}$. Forced air cooling may also be used.

(6) For transistors mounted in a heat sink, of thermal resistance $\theta_{HS\text{-}A}$, the total thermal resistance from junction to ambient is given by

$$\theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{C\text{-}A} \,||\, (\theta_{C\text{-}HS} + \theta_{HS\text{-}A})$$
$$\approx \theta_{J\text{-}C} + \theta_{HS\text{-}A}.$$

(7) If the rate at which heat is released in a transistor is not to exceed the rate at which heat can be dissipated, then

$$V_{CC} - 2I_{CQ}(R_e + R_C)(S)(0.07I_{CO}) < \frac{1}{\theta_{J\text{-}A}}.$$

(8) Thermal fatigue arises from the cyclic stresses that occur in the metal interface between a transistor and the header on which it is mounted. These stresses result from the repeated temperature changes that occur as equipment is turned on and off.

(9) The thermal-cycling rating of a transistor is the number of heating and cooling cycles that a transistor may be subjected to before failure and is a function of case temperature change and power dissipation.

QUESTIONS

Q13-1. What two factors is the power handling capability of a given transistor dependent upon?

Q13-2. What is a power-temperature derating chart? Do all transistors conform to the curves in Fig. 13-1? Can it be used with germanium and silicon transistors alike?

Q13-3. When is a class A CE amplifier inherently stable against thermal runaway? What kind of circuit is generally not inherently stable? Why?

Q13-4. What are the units of measurement for thermal resistance? Explain in your own words what you understand by thermal resistance.

Q13-5. Why is a transistor's junction temperature always higher than the ambient temperature? Which equation supports this?

Q13-6. What two parts make up the junction-to-ambient thermal resistance? Which one is larger?

Q13-7. "The larger the temperature differential between junction and ambient the highe the dissipation capability of the transistor." True or false? Explain.

Q13-8. What is a heat sink? How does it contribute to increasing power dissipation as far as thermal resistance is concerned?

Q13-9. What thermal resistance does an infinite heat sink have? What would a transistor's junction-to-ambient thermal resistance be equal to with an infinite heat sink?

Q13-10. Is a heat sink's thermal resistance constant? Explain.

Q13-11. Under what assumption is $\theta_{J\text{-}A} = \theta_{J\text{-}C} + \theta_{HS\text{-}A}$? What type of mounting would make this equation most accurate, assuming that insulation is necessary between transistor and chassis?

Q13-12. Would the case-to-heat-sink thermal resistance be an important factor in determining the dissipation capability of a signal type transistor? Why or why not?

Q13-13. Under what conditions can a transistor having a given dissipation at 25° C case temperature safely dissipate this power when operated at room temperature of 25° C? Is there any difference between case and ambient temperatures under these conditions?

Q13-14. If a transistor's dissipation at quiescence in a class A amplifier is just below the rated dissipation of the transistor, does this mean the transistor is automatically large enough to avoid overheating? Explain what factors are involved.

Q13-15. Explain the following terms:

(a) thermal fatigue.
(b) thermal cycling.
(c) thermal-cycling rating.

PROBLEMS

P13-1. A 2N5070 silicon RF power transistor has a rated dissipation of 70 W at case temperatures up to 25° C. If its maximum junction operating temperature is 200° C, determine its dissipation capability at the following case temperatures: (a) 50° C, (b) 100°C, (c) 175°C. This transistor conforms to the derating curves in Fig. 13-1.

P13-2. Repeat P13-1 using a 2N5578 silicon power transistor having a rated dissipation of 300 W at T_C up to 25°C. Maximum T_J = 175°C. (Draw a curve for T_J = 175° C.)

P13-3. A 2N649 germanium transistor has a rated dissipation of 100 mW at ambient temperatures up to 25°C. If its maximum junction operating temperature is 71°C, determine its dissipation capability at the following ambient temperatures: (a) 50°C, (b) 65°C, (c) 71° C. Assume ambient and case temperatures to be equal and use the curves in Fig. 13-1.

P13-4. What is the junction-to-ambient thermal resistance of a transistor that has a maximum junction operating temperature of 200° C and can safely dissipate 5 W at an ambient temperature of 25° C?

P13-5. What is the junction-to-case thermal resistance of a silicon 2N5578 transistor that has a maximum junction operating temperature of 175° C and can safely dissipate 300 W at case temperatures up to 25°C?

P13-6. A germanium 2N591 transistor has $\theta_{J\text{-}A}$ = 353° C/W and $\theta_{J\text{-}C}$ = 150° C/W. What is its junction temperature when:

(a) dissipating 50 mW at an ambient temperature of 55° C?
(b) dissipating 100 mW at a case temperature of 55° C?

P13-7. A silicon 2N1701 transistor has $\theta_{J\text{-}A}$ = 100° C/W and $\theta_{J\text{-}C}$ = 7° C/W, with a maximum junction operating temperature of 200° C. Calculate:

(a) The power dissipation capability when operating in an ambient temperature of 25° C.
(b) The power dissipation capability when operating at a case temperature of 50° C.
(c) The rated power dissipation of the 2N1701 at a case temperature of 25° C.
(d) The maximum permissible temperature of the case of the 2N1701 when it is dissipating 20 W.

P13-8. Use Eq.s 13-5 and 13-6 to eliminate T_J by subtraction.(1) Thus obtain an expression for T_C in terms of the other variables and constants. Hence determine:

(a) The temperature of the case of a 2N1701 when working without a heat sink in an ambient temperature of 25° C and dissipating 1 W of power. (Use constants from P13-7.)
(b) The junction temperature for the condition in (a).
(c) Repeat (a) for 2 W of power.
(d) Repeat (b) for 2 W of power.

(2) Rearrange the equation in (1) to obtain an expression for T_A in terms of the other variables and constants. Hence determine:

(a) The maximum ambient temperature if the 2N1701 transistor is working without a heat sink, its case temperature is 120° C, and it is dissipating 1 W of power.
(b) Repeat (a) for 2 W of power, T_C = 186°C.
(c) The maximum ambient temperature if the transistor is working without a heat sink, its case temperature is 25° C, and it is dissipating its rated power of 25 watts. What do you conclude regarding the necessity of a heat sink?

P13-9. A 2N718A is a silicon transistor with $\theta_{J\text{-}C}$ = 97° C/W and $\theta_{J\text{-}A}$ = 350° C/W. It has a maximum dissipation of 0.5 W at ambient temperatures up to 25° C and 1.8 W at case temperatures up to 25° C. It obeys the derating curves of Fig. 13-1 for temperatures above 25° C, and its T_J = 200° C maximum

(a) What is this transistor's $\theta_{C\text{-}A}$?
(b) How much power can the transistor safely dissipate when operating without a heat sink at an ambient temperature of 50° C?
(c) If the transistor is fitted with a fin-type of heat sink having a thermal resistance of 50° C/W, how much power can the 2N718A handle at an ambient temperature of 50°C?
(d) What will be the case temperature for the conditions in (c)? (Solve two ways, including the use of the curves in Fig. 13-1.)

P13-10. A silicon 2N5297 (in a TO-66 package) has a rated dissipation of 36 W at case temperatures up to 25°C. It is mounted in a heat sink ($\theta_{C\text{-}HS}$ = 0) having the characteristics shown in Fig. 13-8. Given:

$T_{J,max}$ = 150° C, $\theta_{J\text{-}C}$ = 3.5° C/W, $\theta_{J\text{-}A}$ =

= 70° C/W, determine:

(a) The power dissipation capability of the transistor and heat sink when operating in free air at an ambient temperature of 25°C.
(b) The temperature of the transistor and heat sink when dissipating the power calculated in (a). Assume $\theta_{C\text{-}HS}$ = 0.
(c) The dissipation capability of the transistor when air at 25°C is forced over the heat sink fins at 0.9 lb/min.
(d) The temperature of the transistor case and heat sink when dissipating the power calculated in (c).

P13-11. Calculate the dissipation capability of the 2N5297 transistor operating under forced air cooling conditions on a heat sink as in P13-10(c) with the following insulators used:

(a) A mica washer with silicone lubricant.
(b) A dry mica washer.

(Use data in Fig. 13-9.)

P13-12. A power transistor is to be mounted on a chassis to provide a heat sink function. If the transistor must dissipate 20 W in an ambient temperature of 50° C, calculate:

(a) The maximum permissible $\theta_{J\text{-}A}$ that the transistor-chassis may have if maximum T_J = 200° C.
(b) The maximum heat-sink-to-air thermal resistance if the transistor's $\theta_{J\text{-}C}$ = 3° C/W, and silicone lubricant is used to make $\theta_{C\text{-}HS}$ = 0.5° C/W with an insulating washer.
(c) The minimum area of chassis in square inches that must be provided if:

(1) The chassis is $\frac{3}{32}$ " thick aluminum and is horizontally mounted.
(2) The chassis is $\frac{1}{16}$ " thick copper and is vertically mounted.

(Use Fig. 13-10.)

P13-13. Consider the amplifier in Fig. 13-11. R_B = 3 K ohms, R_e = 1 ohm, R_C = 6 ohms,

$V_{CC} = 30$ V. If a silicon transistor is used with $\beta = 180$ and $I_{CO} = 0.2$ mA, determine:

(a) If the circuit is inherently stable against thermal runaway. ($V_{BE} = 0.6$ V)

(b) The size of heat sink to use, if necessary, to make the bias point thermally stable. Given $\theta_{J\text{-}C} = 3.5$°C/W and $\theta_{J\text{-}A} =$ 70°C/W.

(c) The power dissipated in the transistor at quiescence.

(d) The maximum junction-to-ambient thermal resistance the transistor may have to dissipate safely the power in (c) if T_J maximum = 175°C and the ambient temperature is 25°C.

(e) The thermal resistance of a heat sink to provide the value of $\theta_{J\text{-}A}$ in (d). Assume $\theta_{C\text{-}HS} = 0$.

(f) The maximum case temperature the transistor can have at quiescence.

(g) The minimum power dissipation capability of the transistor at a case temperature of 25°C. (Use curves from Fig. 13-1.)

P13-14. Repeat P13-13(a) and (b) for a germanium transistor with the same constants except $I_{CO} = 1$ mA, and $V_{BE} = 0.2$ V.

P13-15. A solenoid driver in the peripheral equipment of a computer employs a transistor which has a power dissipation of 15 W and $\theta_{J\text{-}C} = 1.5$°C/W.

(a) If the transistor requires a thermal-cycling rating of 40 million cycles, what must its case temperature change be restricted to in order to avoid failure due to thermal fatigue? (Use Fig. 13-13.)

(b) If an incorrect heat sink were used for the transistor in (a), so that the case temperature change were 50°C, what would be the expected length of life of the transistor expressed in cycles?

(c) Determine the junction temperature of the transistor when operating at an ambient temperature of 25°C for the conditions in (a).

(d) Obtain from (a) and (c) the maximum $\theta_{J\text{-}A}$ the transistor (with heat sink) may have.

(e) If the transistor has a $\theta_{J\text{-}A}$ of 50°C/W, determine the thermal resistance of a heat sink to restrict the case temperature change to the value in (a).

14

HYBRID EQUIVALENT CIRCUIT FOR A TRANSISTOR

14-0. INTRODUCTION

The analysis of transistor amplifiers has been made this far using a graphical approach. However, this gave no real indication of what optimum value to use for R_L in each circuit or how the gains and impedances varied with R_L and R_s. In order to gain a better appreciation of the differences of each amplifier, an equivalent circuit will be developed that is good for small-signal, low-frequency applications. The equivalent circuit and equations assume that the operating point has been chosen correctly and does not involve biasing resistors. However, a practical circuit can always be put in the necessary form so that the equivalent circuit and equations can be used for a given transistor.

14-1. CONVERSION OF AN AMPLIFIER CIRCUIT TO A STANDARD FORM

Consider a practical common-base circuit as shown in Fig. 14-1 (a).

V_{EE}, V_{CC}, R_E, and R_L have been chosen for a suitable operating point. Using Thevenin's theorem, V_s, R_s, and R_E can be replaced by a simple series circuit as far as ac is concerned, since V_{EE} (and V_{CC}) are short circuits for the ac signal.

$$V_g = V_s \times \frac{R_E}{R_s + R_E} = 0.25 \times \frac{1.5\text{ K}}{1\text{ K} + 1.5\text{ K}}$$

$$= 0.25 \times \frac{1.5}{2.5} = 0.15\text{ V rms.}$$

$$R_g = \frac{R_E \times R_s}{R_E + R_s} = \frac{1.5\text{ K} \times 1\text{ K}}{1.5\text{ K} + 1\text{ K}} = \frac{1.5\text{ K}}{2.5}$$

$$= 0.6\text{ K} = 600\text{ ohms.}$$

These are shown in Fig. 14-1 (b), where I_1 and I_2, V_1 and V_2 all represent ac rms values. It is evident that either I_1 or I_2 is negative and flows in the opposite direction. However, since either a PNP or an NPN transistor can be used, this can be taken care of in the resulting equations.

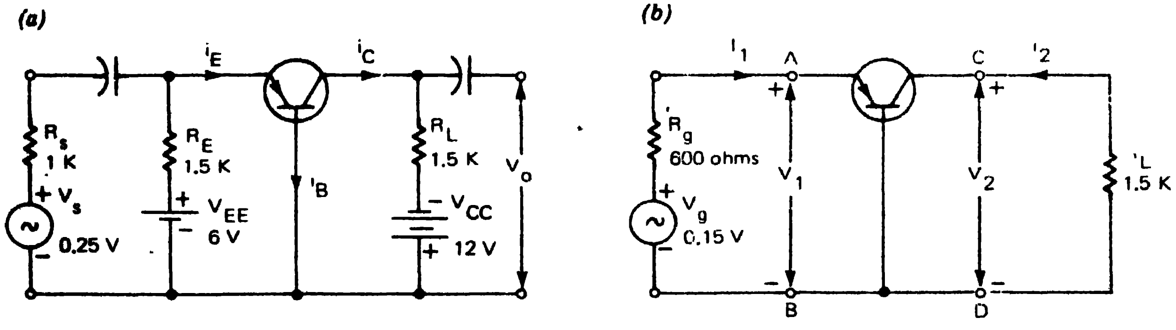


Fig. 14-1. (a) Common-base amplifier circuit complete with bias resistors and dc supplies. (b) Standard form of the amplifier for ac signals, showing V_g, R_g and R_L.

It can be seen that *for a given transistor* the only variables in the circuit of Fig. 14-1 (b) are R_g and R_L. If the section between A-B and C-D is replaced by an equivalent circuit whose values depend upon the transistor in use, we can then determine what effect different values of R_g and R_L will have upon circuit operation.

14-2. GENERAL "BLACK BOX" THEORY

Any four-terminal device (two-terminal input, two-terminal output), no matter how complex, can be represented by an equivalent circuit for small signals provided two restrictions are met:

(1) There must be a common connection between input and output.

(2) The black box must contain only linear elements. The first is met in any of the CE, CB, and CC amplifiers. However, it has been shown that one simple equivalent circuit is one containing diodes, so how can (2) be satisfied? If the signals considered are small (not driving the transistor into saturation or cutoff), then the transistor will be sufficiently linear for the following theory.

Consider a circuit that meets the above restrictions, as represented by the circuit of Fig. 14-2 (a).

If a voltage is applied at the *output* V_2, and I_2 allowed to flow, the Thevenin equivalent circuit looking into the *input* terminals is given by a voltage source V_{01} in series with an impedance Z_{01}. Similarly, if a voltage V_1 is applied at the *input* , the Thevenin equivalent circuit looking back from the *output* terminals is a voltage source V_{02} in series with an impedance Z_{02}. These are shown in Fig. 14-2 (b) and can be found for any device, transistor, transformer, etc.

However, with a transistor, it is convenient to have a slightly different equivalent circuit in the output, called a Norton equivalent instead of a Thevenin equivalent. This consists of a current generator in parallel with an impedance. The impedance is Z_{02} and the current generator has a magnitude equal to

$\frac{V_{02}}{Z_{02}}$ with a direction as shown in Fig. 14-3.

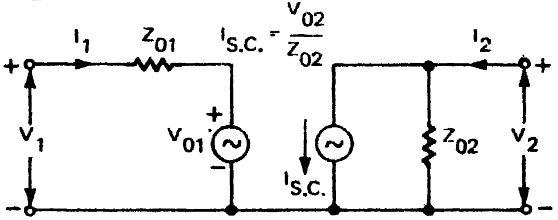


Fig. 14-3. Transistor equivalent circuit showing a Thevenin voltage equivalent for the input and a Norton current equivalent for the output.

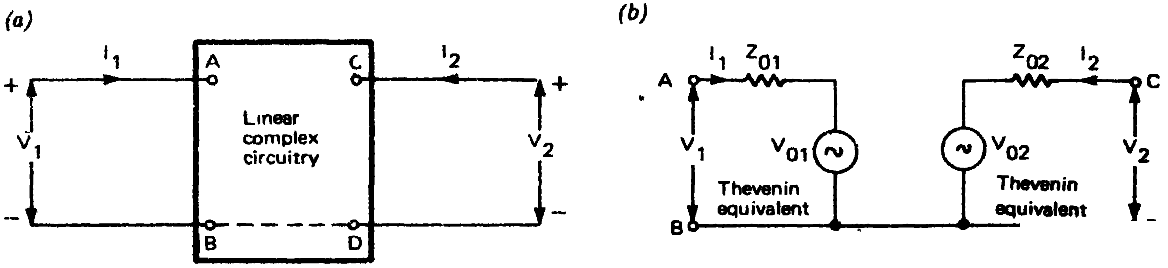


Fig. 14-2. (a) General "black box" showing current and voltage polarities.
(b) Thevenin equivalent circuit for the linear "black box."

14-3. HYBRID h-PARAMETERS

Our next task is to determine the relations between these variables and constants. Since the diagram in Fig. 14-3 is applicable to other devices as well as the transistor, we will use general symbols now, but use specific ones for transistors later.

(a) If applying a voltage V_1 to a device causes a current I_1 to flow, then we can calculate the input resistance. Accordingly, Z_{01} is given the symbol h_{11} and represents the input resistance.

(b) It has been shown that the input characteristics of a transistor depend upon what value of voltage exists at the output. Thus V_{01} is related to V_2 by some constant of proportionality equal to h_{12} such that $V_{01} = h_{12}\ V_2$.

(c) The current that flows in the output is dependent upon the input current. Hence the current generator $I = h_{21}\ I_1$.

(d) Finally, the output of a device has some output impedance or resistance. With the transistor it is more useful to use a conductance instead of a resistance. Therefore, in place of Z_{02}, we use a conductance of value h_{22}.

The resulting hybrid equivalent circuit is shown in Fig. 14-4. (Hybrid because it is a mixture of constants having different units of ohms, mhos, volts and amps.)

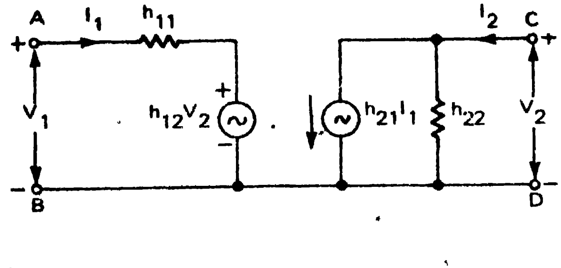


Fig. 14-4. Hybrid h-parameter equivalent circuit for a linear four-terminal network.

The equations relating input and output variables are:

$$V_1 = h_{11}I_1 + h_{12}V_2 \tag{14-1}$$

$$I_2 = h_{21}I_1 + h_{22}V_2. \tag{14-2}$$

Since V_1, V_2, I_1, and I_2 are the only variables from the original amplifier circuit, the transistor has been replaced by four h-parameter constants h_{11}, h_{12}, h_{21}, and h_{22}. They are constant for a given transistor (although they vary somewhat with collector current), but will be different in value for each circuit the transistor is used in, CE, CB, or CC.

Accordingly, it is wise to use a subscript notation which will identify the parameter as to whether it can be used in a CE, CB, or CC circuit.

14-4. OBTAINING THE HYBRID h-PARAMETERS

Let us consider this new subscript notation along with how these four parameters (constants) can be determined for a CE connection.

The Parameter h_{11} (h_{ie})

Consider the equation: $V_1 = h_{11}I_1 + h_{12}V_2$

If V_2 is made zero, then $V_1 = h_{11}I_1$, and

$$h_{11} = \frac{V_1}{I_1}$$

This is evidently the input resistance to the transistor with the output shorted to ac signals, and may be obtained by applying a voltage at the input V_1 and measuring I_1.

$$\text{Thus } h_{11} = \text{input resistance} = \frac{V_1}{I_1},\ (V_2 = 0)$$

$$= h_{ib},\ h_{ie},\ \text{or } h_{ic}$$

depending upon whether a CB, CE, or CC connection is being used for the transistor. h_{11} is measured in ohms.

Note the notation:

h_{ib} means *input* resistance in common *base*.
h_{ie} means *input* resistance in common *emitter*.
h_{ic} means *input* resistance in common *collector*.

Remember that V_1 and I_1 are the rms values of the applied ac signal, usually at a test frequency of 1 kHz. One method of obtaining h_{ie} is to use the CE input characteristic curves for the transistor, relating i_B to v_B. Then:

$$h_{ie} \approx \frac{\Delta v_B}{\Delta i_B},\ V_C \text{ constant,} \tag{14-3}$$

and is the reciprocal of the slope of the tangent drawn to the i_B-v_B curve in the region of the operating point, for a fixed value of V_C, as shown in Fig. 14-5 (a).

Alternatively, h_{ie} could be obtained without the use of ac voltages, by causing a change in base voltage level, Δv_B and noting the change in base current level, Δi_B. These should be small changes and made to occur in the region of the operating point.

The Parameter h_{21} (h_{fe})

If, while V_1 was applied, I_2 was measured as well as I_1, then, in the equation, $I_2 = h_{21}I_1 + h_{22}V_2$, if $V_2 = 0$, $I_2 = h_{21}I_1$, and

$$h_{21} = \frac{I_2}{I_1}.$$

This we have seen before as the ac forward current transfer ratio, α_{ac} in a CB, β_{ac} in a CE. However, since our equivalent circuit has I_1 and I_2 both going *into* the transistor, h_{21} will be negative for a CB and CC when using this circuit.

Thus, h_{21} = forward current transfer ratio

$$= \frac{I_2}{I_1},\ (V_2 = 0) = h_{fb}, h_{fe}, \text{ or } h_{fc},$$

depending upon whether the transistor is used in a CB, CE, or CC amplifier.

Note: h_{fb} means *forward* current transfer ratio in common *base*, etc., and is a pure number, being the ratio of amps to amps.

Fig. 14-5 (b) shows how the output characteristic curves may be used to evaluate h_{fe} in the region of the operating point at a fixed collector voltage. That is,

$$h_{fe} \approx \frac{\Delta i_C}{\Delta i_B},\ V_C \text{ constant, } = \frac{i_{C2} - i_{C1}}{i_{B2} - i_{B1}}. \tag{14-4}$$

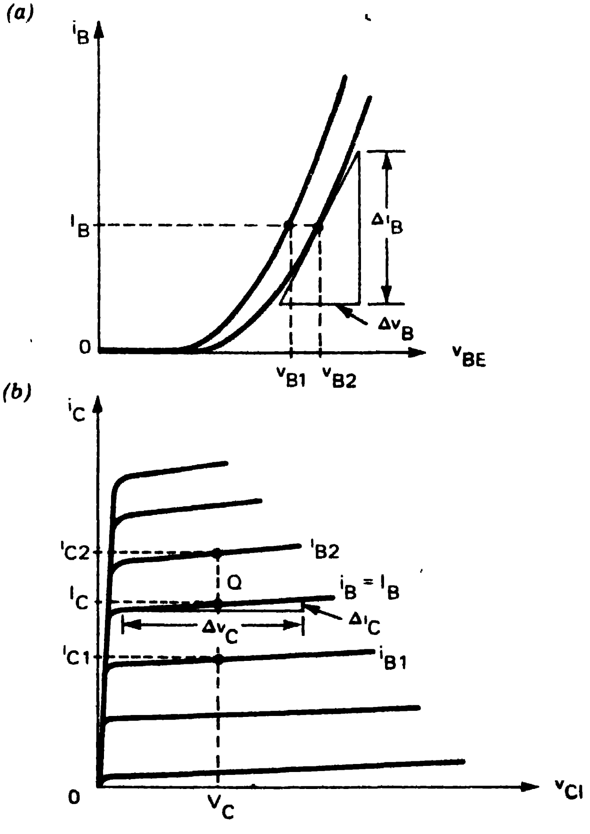


Fig. 14-5. (a) Showing how the CE input characteristic curves are used to determine h_{ie} and h_{re}.
(b) Showing how the CE output characteristic curves are used to determine h_{fe} and h_{oe}.

And, of course, this may be evaluated by using the ratio of the change in collector current level to the change in base current level, instead of applying an ac signal.

The Parameter h_{12} (h_{re})

Referring to $V_1 = h_{11}I_1 + h_{12}V_2$, if $I_1 = 0$, then $V_1 = h_{12}V_2$, and $h_{12} = \frac{V_1}{V_2}$.

This can be measured by applying a voltage V_2 at the output of the transistor with the input open ($I_1 = 0$) and measuring the voltage that appears at the input, V_1 (working the transistor in reverse). This is called the reverse voltage transfer ratio.

Thus h_{12} = reverse voltage transfer ratio

$$= \frac{V_1}{V_2}, (I_1 = 0) = h_{rb}, h_{re}, h_{rc}.$$

Note: h_{rb} means *reverse* voltage ratio in a common *base*, etc., and is a dimensionless number.

It should be recalled from Chapter 9 that the CE input characteristic for a transistor was dependent upon the output voltage. See Fig. 14-5 (a). At some fixed value of I_B, a change in the necessary input voltage from v_{B1} to v_{B2} is caused by a change in the output voltage from v_{C1} to v_{C2}. Thus:

$$h_{re} \approx \frac{\Delta v_B}{\Delta v_C}, I_B \text{ constant.} \qquad (14\text{-}5)$$

h_{re} is typically in the order of 1×10^{-4}. That is, $\Delta v_B << \Delta v_C$, so the method of using a change in levels is not very accurate.

The Parameter h_{22} (h_{oe})

If, while applying V_2, the current I_2 is measured with the input open ($I_1 = 0$), then $I_2 = h_{22}V_2$, and $h_{22} = I_2/V_2$.

This is obviously a measure of the output conductance in mhos, with the input open-circuited.

Thus, h_{22} = output conductance

$$= \frac{I_2}{V_2}, (I_1 = 0) = h_{ob}, h_{oe}, \text{or } h_{oc}$$

Note: h_{ob} means *output* conductance in a common *base*, etc.

The ratio of $\frac{I_2}{V_2}$ for a common emitter determines h_{oe} and may be seen from Fig. 14-5 (b) to correspond to the ratio of the change in collector current to the change in collector voltage for a constant base current. That is,

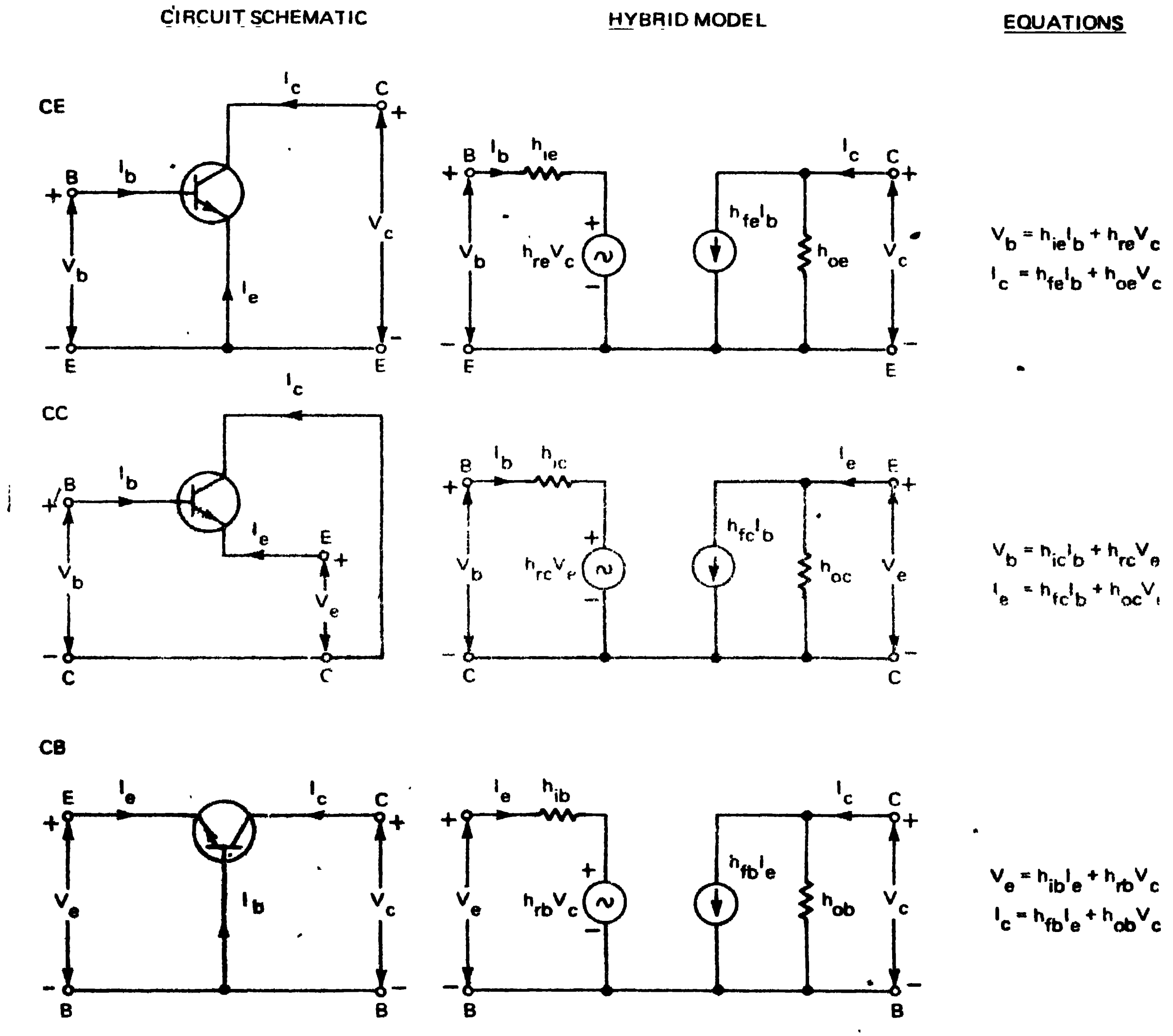
$$h_{oe} \approx \frac{\Delta i_C}{\Delta v_C}, I_B \text{ constant.} \qquad (14\text{-}6)$$

This is the slope of the output characteristic curves in the region of the operating point, and may easily be measured by using a change in voltage and current levels.

The method of obtaining the h-parameters for a transistor in the CB or CC configuration is the same as that described above for the CE using the appropriate characteristic curves or changes in voltage and current levels.

14-5. TYPICAL h-PARAMETER VALUES

No matter whether a CB, CE, or CC amplifier is being considered, the same equivalent circuit configuration can be used as long as the appropriate numerical values for the h-parameters are inserted. Fig. 14-6 gives the circuit schematic, hybrid model, and V-I equations for the CE, CC, and CB amplifier connections.


CIRCUIT SCHEMATIC
HYBRID MODEL
EQUATIONS
CE
CC
CB
B
C
E
I_b
I_c
I_e
V_b
V_c
V_e
h_{ie}
$h_{re}V_c$
$h_{fe}I_b$
h_{oe}
h_{ic}
$h_{rc}V_e$
$h_{fc}I_b$
h_{oc}
h_{ib}
$h_{rb}V_c$
$h_{fb}I_e$
h_{ob}

$$V_b = h_{ie}I_b + h_{re}V_c$$
$$I_c = h_{fe}I_b + h_{oe}V_c$$

$$V_b = h_{ic}I_b + h_{rc}V_e$$
$$I_e = h_{fc}I_b + h_{oc}V_e$$

$$V_e = h_{ib}I_e + h_{rb}V_c$$
$$I_c = h_{fb}I_e + h_{ob}V_c$$

Fig. 14-6. Circuit schematic, hybrid model, and V-I h-parameter equations for the CE, CC, and CB amplifier connections.

These equations and equivalent circuits hold for PNP, NPN, silicon, or germanium transistors, at low frequencies.

Table 14-1 shows the typical values for a TIS98 (2N3391) silicon transistor's h-parameters evaluated at an operating point

TABLE 14-1 TYPICAL h-PARAMETERS FOR A 2N3391 SILICON TRANSISTOR, EVALUATED AT $I_C = 1$ mA, $V_{CE} = 5$ V, f = 1 kHz, $T_A = 25°C$

PARAMETER	CE	CC	CB
$h_{11} = h_i$	6,400 ohms	6,400 ohms	26.6 ohms
$h_{12} = h_r$	1.5×10^{-4}	1	0.1×10^{-4}
$h_{21} = h_f$	240	241	-0.995
$h_{22} = h_o$	6 μmhos	6 μmhos	0.025 μmhos
$1/h_o$	166 K ohms	166 K ohms	40 M ohms

given by $I_C = 1$ mA, $V_{CE} = 5$ V, and at a frequency of 1 kHz and 25°C ambient temperature.

These parameters are dependent upon the collector current as shown by the normalized curves in Fig. 14-7. For example, at a collector current of $I_C = 10$ mA, the CE h-parameters become:

$h_{ie} = 1.28$ K ohms, $h_{re} = 0.8 \times 10^{-4}$, $h_{fe} = 250$, $h_{oe} = 42$ μmhos.

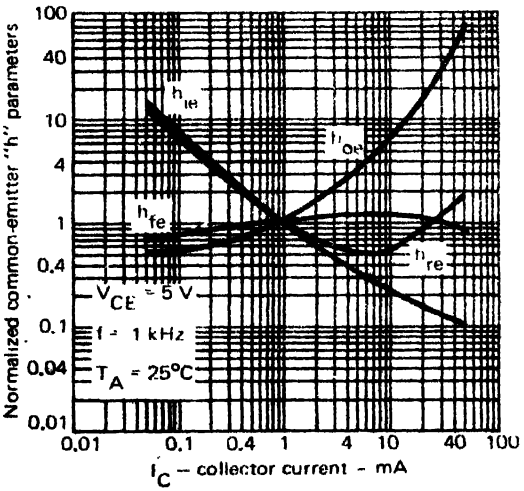


Fig. 14-7. Normalized curves showing the variation of h-parameter values with collector current in a 2N3391 transistor. (Courtesy, Texas Instruments Incorporated.)

Typical h-parameters for a germanium 2N104 transistor are shown in Table 14-2 and were evaluated at $I_C = 1$ mA, $V_{CE} = 6$ V, f = 1 kHz, and $T_A = 25°C$. These also are dependent upon the value of I_C. Note that h_{fe} is also dependent upon temperature, as mentioned in Chapter 9.

Tables 14-1 and 14-2 include values for $1/h_o$. Note how high the output resistance is in a CB compared with a CE or CC connection.

At one time, some transistor manufacturers only gave the h-parameters for the CB connection. Now, most information provides the CE h-parameters. For some transistors, however, a mixture of CB and CE parameters is used; for example, in the 2N1302:

$h_{ib} = 28$ ohms, $h_{rb} = 5 \times 10^{-4}$, $h_{ob} = 0.34$ μmhos, $h_{fe} = 105$. If the transistor is to be used in a connection different from the one for which the h-parameters were specified conversion formulas may be used. Table 14-3 provides the approximate conversion formulas to obtain any of the parameters from the given CE or CB values.

For example, if h_{fb} is given as - 0.98 then h_{fe} is found as follows:

$$h_{fe} = \frac{-h_{fb}}{1 + h_{fb}} = \frac{-(-0.98)}{1 - 0.98} = \frac{0.98}{0.02} = 49.$$

TABLE 14-2 TYPICAL h-PARAMETERS FOR A 2N404 GERMANIUM TRANSISTOR, EVALUATED AT $I_C = 1$ mA, $V_{CE} = 6$ V, f = 1 kHz, $T_A = 25^oC$

PARAMETER	CE	CC	CB
$h_{11} = h_i$	4,000 ohms	4,000 ohms	29.4 ohms
$h_{12} = h_r$	7×10^{-4}	1	7.7×10^{-4}
$h_{21} = h_f$	135	- 136	- 0.993
$h_{22} = h_o$	50 μmhos	50 μmhos	0.37 μmhos
$1/h_o$	20 K ohms	20 K ohms	2.7 M ohms

TABLE 14-3 APPROXIMATE CONVERSION FORMULAS FOR TRANSISTOR h-PARAMETERS

PARAMETER	CE	CB
h_{ie}	—	$\frac{h_{ib}}{1 + h_{fb}}$
h_{re}	—	$\frac{h_{ib}h_{ob}}{1 + h_{fb}} - h_{rb}$
h_{fe}	—	$\frac{-h_{fb}}{1 + h_{fb}}$
h_{oe}	—	$\frac{h_{ob}}{1 + h_{fb}}$
h_{ib}	$\frac{h_{ie}}{1 + h_{fe}}$	—
h_{rb}	$\frac{h_{ie}h_{oe}}{1 + h_{fe}} - h_{re}$	—
h_{fb}	$\frac{-h_{fe}}{1 + h_{fe}}$	—
h_{ob}	$\frac{h_{oe}}{1 + h_{fe}}$	—
h_{ic}	h_{ie}	$\frac{h_{ib}}{1 + h_{fb}}$
h_{rc}	$1 - h_{re} \approx 1$	1
h_{fc}	$-(1 + h_{fe})$	$\frac{-1}{1 + h_{fb}}$
h_{oc}	h_{oe}	$\frac{h_{ob}}{1 + h_{fb}}$

14-6. AMPLIFIER EQUATIONS

Let it be required to find the equations for A_i, A_v, A_p, R_i, and R_o for a CB, CE, or CC amplifier in terms of the h-parameters and R_g, R_L. See Fig. 14-8.

Since the hybrid equivalent circuit is applicable to CB, CE, or CC, the h-parameters will be referred to as h_i, h_r, h_f, h_o. The resulting equations will then apply to whichever circuit is being considered simply by putting in the appropriate second subscript.

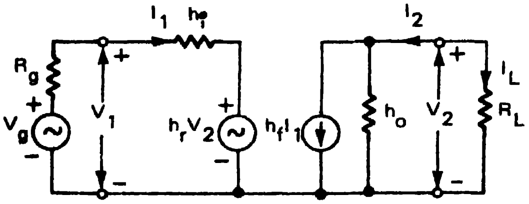


Fig. 14-8. Hybrid equivalent circuit for a CB, CE, or CC amplifier.

Current Gain

In Chapter 7, we defined the current gain to be the ratio of output current to input current. Thus

$$A_i = \frac{I_L}{I_1} = \frac{-I_2}{I_1}.$$

But $I_2 = h_fI_1 + h_oV_2$, and $V_2 = I_LR_L = -I_2R_L$.

$$\therefore I_2 = h_fI_1 - h_oI_2R_L$$

$$I_2 + h_oR_LI_2 = h_fI_1$$

$$I_2(1 + h_oR_L) = h_fI_1.$$

$$\therefore \frac{I_2}{I_1} = \frac{h_f}{1 + h_oR_L}, \text{ and}$$

$$A_i = \frac{-I_2}{I_1} = \frac{-h_f}{1 + h_oR_L}. \tag{14-7}$$

If a CE circuit is being considered, then A_i would be

$$A_i = \frac{-h_{fe}}{1 + h_{oe}R_L},$$

and with h_{fe} a positive number, say 50, A_i would be negative.

This negative value for A_i must be interpreted with respect to the directions of I_L and I_1, shown in Fig. 14-8. For example in an NPN CE amplifier, where base current I_1 flows *into* the base, the output load current I_c flows *into* the collector and is *opposite* in direction to that assumed for I_L in Fig. 14-8. This is the meaning of the negative sign. In a PNP CE connection, I_L *is* in the direction assumed, but I_1 is opposite to that assumed in Fig. 14-8. (Sketch the CE transistor amplifier and verify this.)

If a CB circuit is being considered, then A_i would be

$$A_i = \frac{-h_{fb}}{1 + h_{ob}R_L},$$

and with h_{fb} a negative number, say - 0.98, A_i would be positive. This would show that both the input and output currents (I_e and I_c) have the same directions as I_1 and I_L in Fig. 14-8 for a PNP transistor, or *both* currents are opposite in an NPN transistor.

Input Resistance

$$R_i = \frac{V_1}{I_1}.$$

But $V_1 = h_iI_1 + h_rV_2$,

and $V_2 = -I_2R_L$.

$$\therefore V_1 = h_iI_1 - h_rI_2R_L,$$

and $\frac{V_1}{I_1} = \frac{h_iI_1 - h_rI_2R_L}{I_1} = h_i - h_rR_L\frac{I_2}{I_1}$.

But $\frac{I_2}{I_1} = -A_i$.

$$\therefore R_i = \frac{V_1}{I_1} = h_i + h_rR_LA_i = h_i + h_rA_iR_L. \tag{14-8}$$

Note that the input resistance depends upon the load resistance.

If the input resistance of a CE circuit is being calculated, then

$R_i = h_{ie} + h_{re}A_iR_L$,

and, since h_{re} is positive and A_i is negative in a CE, then the input resistance, R_i, is a little less than h_{ie}. In a CB circuit, however, A_i is positive, so that R_i is slightly in excess of h_{ib}.

Voltage Gain

$A_v = \frac{V_2}{V_1}$.

But $V_2 = -I_2R_L$, and $I_2 = -A_iI_1$.

$\therefore V_2 = A_iI_1R_L = A_iR_LI_1$.

$A_v = \frac{V_2}{V_1} = \frac{A_iR_LI_1}{V_1} = A_iR_L \times \frac{I_1}{V_1}$.

But $\frac{V_1}{I_1} = R_i$. $\therefore \frac{I_1}{V_1} = \frac{1}{R_i}$

$$\therefore A_v = \frac{A_iR_L}{R_i} \tag{14-9}$$

This could be put in terms of the h-parameters, but the resulting equation is complicated, whereas the above equation shows clearly the dependence of A_v on R_L. Also, in any problem, it is usual to want to know what A_i and R_i are anyway, so A_v follows quickly.

Examining the equation for A_v, R_L and R_i must both be positive, therefore A_v will be negative in a CE amplifier (since A_i is negative) showing phase inversion between input and output. In a CB circuit, A_i is positive, so A_v is positive, showing no phase shift between input and output voltage.

Output Resistance

$R_o = \frac{V_2}{I_2}$, or $G_o = \frac{I_2}{V_2}$ = output conductance.

It can be shown from the preceding equations that

$$G_o = h_o - \frac{h_fh_r}{h_i + R_g}, \tag{14-10}$$

and $R_o = \frac{1}{G_o}$.

Note that the output resistance depends upon the generator or source resistance R_g.

Power Gain

$$A_p = A_v\,A_i. \tag{14-11}$$

Since A_v and A_i are either both positive or both negative, A_p will always be positive.

14-7. VOLTAGE AND CURRENT GAINS TAKING INTO ACCOUNT R_g OF SOURCE

The above gains were calculated from the *input* of the transistor to the output, whereas it is often more meaningful to get the gain from the generator open circuit voltage V_g. Refer to Fig. 14-9.

A_{vg} (voltage gain from V_g to V_2)

$$= \frac{V_2}{V_g} = \frac{A_vR_i}{R_i + R_g}. \tag{14-12}$$

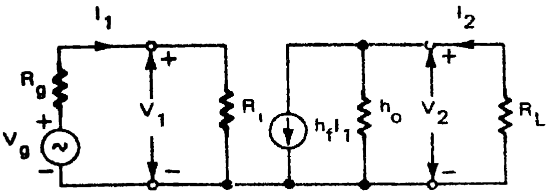


Fig. 14-9. Equivalent circuit for a transistor, showing the input resistance, R_i, and the open circuit voltage of the source, V_g, and resistance, R_g.

Thus, if $R_g = R_i$ for example, the *overall* amplifier voltage gain, A_{vg}, would be one-half of the transistor's voltage gain, A_v. This is similar to the discussion in Section 10-2. To see the difference between the overall current gain, A_{ig}, and the transistor current gain, A_i, consider the Norton equivalent for the source shown in Fig. 14-10.

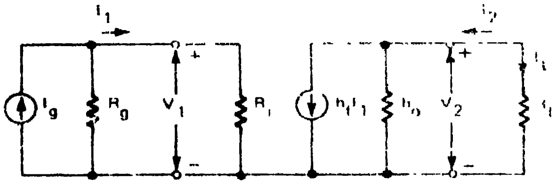


Fig. 14-10. Norton equivalent circuit for a transistor, showing the input resistance, R_i, and the generator constant current source, I_g, and resistance, R_g.

The current I_g will divide up inversely between R_g and R_i, so that the *overall* amplifier's current gain, A_{ig}, is given by

$$A_{ig} = \frac{I_L}{I_g} = \frac{-I_2}{I_g} = \frac{A_i R_g}{R_i + R_g}. \tag{14-13}$$

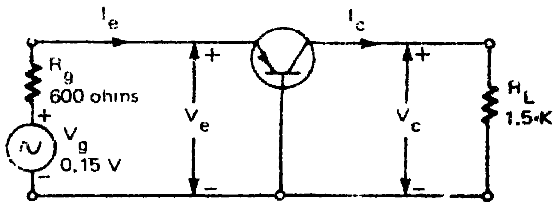
Only if $R_g >> R_i$ does $A_{ig} = A_i$; otherwise A_{ig} is less than A_i.

The equations developed in the last two sections (which are exact and assume no approximations) are summarized in Table 14-4. The second subscript for the h-parameters is determined by whichever circuit the equations are applied to: CB, CE or CC.

TABLE 14-4 EXACT EQUATIONS FOR THE SMALL-SIGNAL ANALYSIS OF TRANSISTOR AMPLIFIERS

$A_i = \dfrac{-h_f}{1 + h_o R_L}$	$G_o = \dfrac{1}{R_o} = h_o - \dfrac{h_f h_r}{h_i + R_g}$
$R_i = h_i + h_r A_i R_L$	$A_{vg} = \dfrac{A_v R_i}{R_i + R_g}$
$A_v = \dfrac{A_i R_L}{R_i}$	$A_{ig} = \dfrac{A_i R_g}{R_i + R_g}$

Example 14-1. As an example of using these equations, consider the CB amplifier we first started with in Section 14-1. Fig. 14-11 shows the circuit in standard form for small-signal analysis, along with the CB h-parameters.



h_{ib} = 20 ohms
h_{rb} = 3×10^{-4}
h_{fb} = -0.98
h_{ob} = 0.5×10^{-6} mho

Fig. 14-11. Standard form of CB amplifier for ac signal analysis, and the CB h-parameters for the transistor.

Let it be required to find A_i, R_i, A_v, R_o, A_p, A_{vg}, and A_{ig}.

Solution

(1) $$A_i = \frac{I_c}{I_e} = \frac{-h_{fb}}{1 + h_{ob}R_L}$$

$$= \frac{-(-0.98)}{1 + 0.5 \times 10^{-6} \times 1.5 \times 10^3}$$

$$= \frac{0.98}{1 + 0.75 \times 10^{-3}} = 0.98.$$

(2) $$R_i = h_{ib} + h_{rb}A_iR_L$$

$$= 20 + 3 \times 10^{-4} \times 0.98 \times 1.5 \times 10^3$$

$$= 20 + 4.41 \times 10^{-1} = 20.44 \text{ ohms}.$$

(3) $$A_v = \frac{A_i \times R_L}{R_i} = \frac{0.98 \times 1.5 \times 10^3}{20.44} = 72.$$

(4) $$G_o = h_{ob} - \frac{h_{fb}h_{rb}}{h_{ib} + R_g}$$

$$= 0.5 \times 10^{-6} - \frac{(-0.98) \times 3 \times 10^{-4}}{20 + 600}$$

$$= 0.5 \times 10^{-6} + \frac{2.94 \times 10^{-4}}{620}$$

$$= 0.5 \times 10^{-6} + 0.47 \times 10^{-6}$$

$$= 0.97 \times 10^{-6} \text{ mhos}.$$

$$R_o = \frac{1}{G_o} = \frac{1}{0.97 \times 10^{-6}} = 1.03 \times 10^6 \text{ ohms}$$

$$= 1.03 \text{ M ohms}.$$

(5) $A_p = A_iA_v = 0.98 \times 72 = 70.56$.
Gain in dB = 10 log A_p = 10 log 70.56
= 10 × 1.85 = 18.5 dB.

This substantiates our previous findings that a CB amplifier has a very low input impedance, a high output impedance, and no phase inversion between input and output.

Note now the large effect of R_g upon the voltage gain:

(6) $$A_{vg} = \frac{V_c}{V_g} = \frac{A_vR_i}{R_i + R_g} = 72 \times \frac{20.44}{20.44 + 600}$$

$$= 72 \times \frac{20.44}{620.44} = 2.3.$$

Thus it would be very impractical to drive this CB amplifier from a source of internal resistance 600 ohms.

If the voltage gain $A_{vs} = \frac{V_c}{V_s}$ were required (see Fig. 14-1 (a)), then:

$$A_{vs} = A_{vg}\frac{V_g}{V_s} = 2.3 \times \frac{0.15}{0.25} = 1.38.$$

(7) $$A_{ig} = \frac{I_c}{I_g} = \frac{A_iR_g}{R_i + R_g} = \frac{0.98 \times 600}{20.44 + 600} = 0.95.$$

14-8. DUAL MATCHING IN AN AMPLIFIER FOR MAXIMUM POWER GAIN

It is possible to determine, since R_i depends upon R_L, and R_o upon R_g, what dual matching condition will yield maximum power gain. For this condition, R_L is equal to the output resistance of the amplifier, R_o, and R_g equals the input resistance, R_i. To satisfy this condition the necessary values of R_L and R_g are given by the following equations:

$$R_L = \sqrt{\frac{h_i}{\Delta h \times h_o}} \tag{14-14}$$

$$R_g = \sqrt{\frac{\Delta h \times h_i}{h_o}} \tag{14-15}$$

where $\Delta h = h_ih_o - h_rh_f$.

However, these are not always possible values to use, so this is more of academic interest, although it does point up the dependence of gain upon R_L and R_g. For example, it is possible for a typical transistor to provide in the CE connection a power gain of over 14,000 (42 dB) when the input and output are dually matched.

Example 14-2. Determine what would be the optimum values of R_g and R_L in Fig. 14-11 to provide maximum power gain using the CB h-parameters provided. Also determine this power gain.

Solution. Using Eqs. 14-14 and 14-15, we need first to evaluate Δh.

$$\Delta h = h_{ib}h_{ob} - h_{rb}h_{fb}$$

$$= 20 \times 0.5 \times 10^{-6} - 3 \times 10^{-4} \times (-0.98)$$

$$= 1 \times 10^{-5} + 29.4 \times 10^{-5}$$

$$= 30.4 \times 10^{-5}.$$

$$R_L = \sqrt{\frac{h_{ib}}{\Delta h \times h_{ob}}} = \sqrt{\frac{20}{30.4 \times 10^{-5} \times 0.5 \times 10^{-6}}}$$

$$= \sqrt{13.2 \times 10^{10}} = 364 \text{ K ohms}.$$

$$R_g = \sqrt{\frac{\Delta h \times h_{ib}}{h_{ob}}} = \sqrt{\frac{30.4 \times 10^{-5} \times 20}{0.5 \times 10^{-6}}}$$

$$= \sqrt{122 \times 10^{2}} = 110 \text{ ohms}.$$

$$A_i = \frac{-h_{fb}}{1 + h_{ob}R_L} = \frac{-(-0.98)}{1 + 0.5 \times 10^{-6} \times 3.64 \times 10^{5}}$$

$$= \frac{0.98}{1 + 0.182} = 0.83.$$

Since the amplifier is dually matched, $R_i = R_g = 110$ ohms.

$$\therefore A_v = \frac{A_i R_L}{R_i} = \frac{0.83 \times 3.64 \times 10^{5}}{110} = 2{,}750.$$

$$A_p = A_v A_i = 2{,}750 \times 0.83 = 2{,}280.$$

$$\text{dB gain} = 10 \log A_p = 10 \log 2{,}280$$
$$= 10 \times 3.36 = 33.6 \text{ dB}.$$

14-9. DEPENDENCE OF AMPLIFIER CHARACTERISTICS ON R_L AND R_g

It should be obvious from a comparison of the results of Example 14-2 with those in Example 14-1 that A_i, A_v, and R_i have a strong dependence on R_L. Also, R_o depends upon the value of R_g. The variation of these characteristics with the load and generator resistances may best be seen by referring to the curves in Fig. 14-12. The values shown for the three amplifiers are typical for a transistor having h-parameters of the values used in Example 14-1.

14-10. COMPARISON OF CB, CE, AND CC CHARACTERISTICS

Although the curves in Fig. 14-12 show the difference in A_i, A_v, R_i, and R_o for any given value of R_L and R_g, the comparison may be made more easily by looking at Table 14-5. This has been developed using the transistor's h-parameters listed in Example 14-1 and for the same values of R_g and R_L for each amplifier as were used in Example 14-1. R_o is the effective output resistance at the overall amplifier's output terminals and consists of R_L in parallel with R_o. R_o' is the significant value as far as any external load connected to the amplifier is concerned.

The determination of the CE and CC amplifier characteristics in Table 14-5 required the use of the conversion formulas in Table 14-3 to first obtain the appropriate

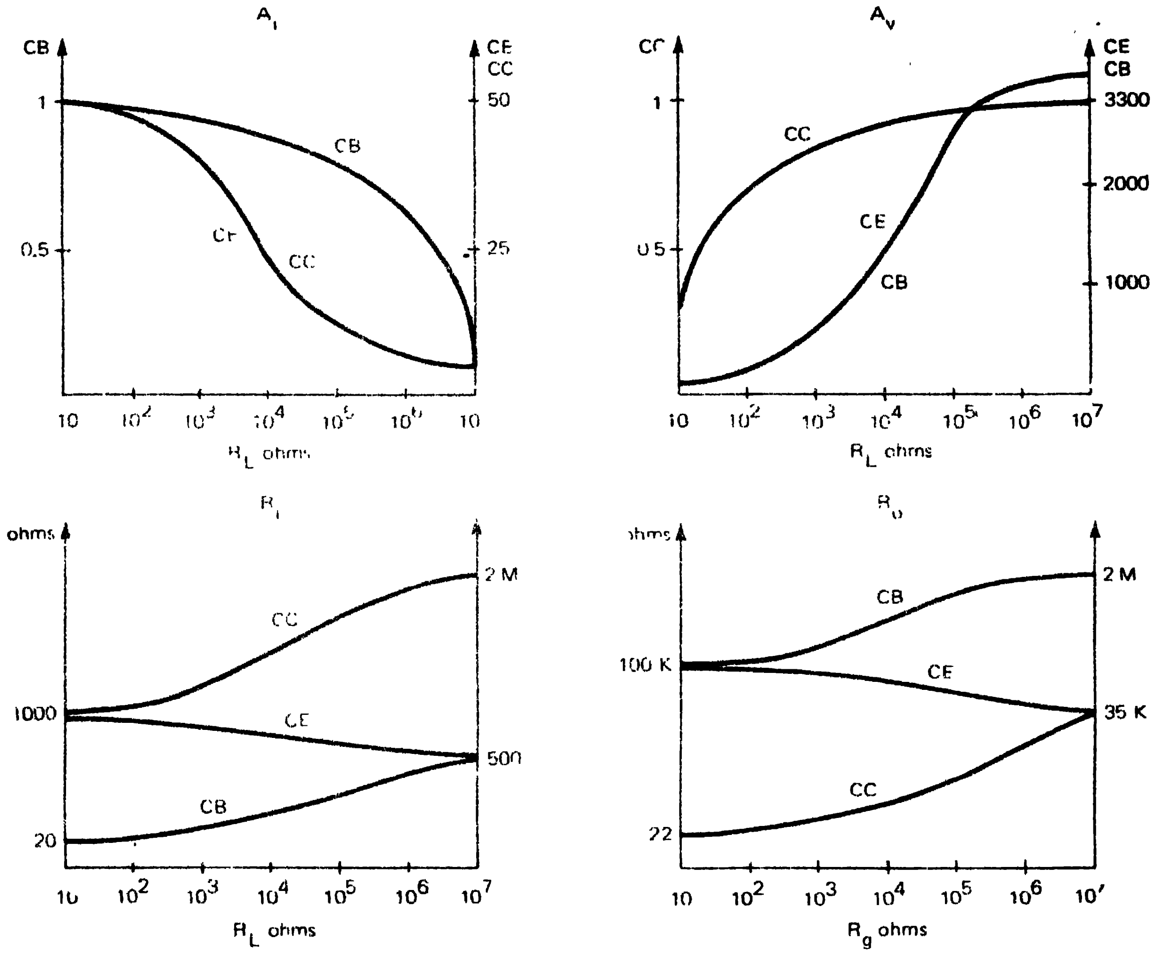


Fig. 14-12. Curves showing the variation of A_i, A_v, R_i, and R_o with R_L and R_g, for the three amplifier connections using the same transistor as in Example 14-1.

h-parameters from the given CB h-parameters. These were then used in the *exact* equations in Table 14-4 to obtain A_i, A_v, R_i, and R_o. Although in this example the given h-parameters were for the CB, it is more usual that the given h-parameters are for the CE, since it is the more popular circuit. It is thus found convenient to express the equations for A_i, A_v, R_i, and R_o in terms of the CE h-parameters, for all three amplifier connections. These equations may be made more simple and within 10% of the exact equations by assuming that $h_{oe}R_L$ (or $h_{oe}R_e$) ≤ 0.1. The equivalent circuit resulting from this approximation is shown in Fig. 14-13. These *approximate* equations are shown in Table 14-6 and include a column for a CE amplifier with an unbypassed resistor, R_e, in the emitter lead.

TABLE 14-5 COMPARISON OF AMPLIFIER CHARACTERISTICS CALCULATED USING THE EXACT EQUATIONS OF TABLE 14-4, WITH $R_g = 600$ OHMS, $R_L = 1.5$ K OHMS, $h_{ib} = 20$ OHMS, $h_{rb} = 3 \times 10^{-4}$ $h_{fb} = -0.98$, $h_{ob} = 0.5 \times 10^{-6}$ MHO

Quantity	CB	CE	CC
A_i	Low (0.98)	High (-47.2)	High (48.2)
A_v	High (72)	High (-72)	Low (0.99)
R_i, ohms	Low (20.4)	Med. (986)	High (73 K)
R_o, ohms	High (1.03 M)	Med. (52 6 K)	Low (32)
$R_o'(R_L \| R_o)$, ohms	1.5 K	1.46 K	31

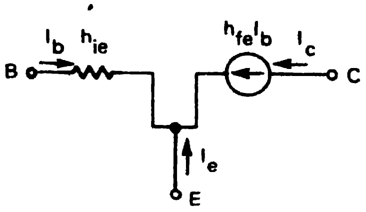


Fig. 14-13. Approximate hybrid equivalent circuit at low frequencies, when $h_{oe}R_L \leq 0.1$. Note: $h_{oe} = h_{re} = 0$.

To illustrate how close these approximate equations are to the exact equations, A_i, A_v, R_i, and R_o have been calculated using the same h-parameters and conditions for R_L and R_g as for the amplifiers in Table 14-5, where the exact equations were used. These approximate characteristics are listed in Table 14-7 for comparison purposes.

Note: $h_{oe}R_L = 25 \times 10^{-6} \times 1.5 \times 10^3 = 37.5 \times 10^{-3} \approx 0.038 < 0.1$.

TABLE 14-6 APPROXIMATE EQUATIONS FOR A_i, R_i, A_v, AND R_o FOR CB, CE, OR CC AMPLIFIERS, GIVEN IN TERMS OF CE h-PARAMETERS: TO BE USED ONLY WHEN $h_{oe}R_L$ OR $h_{oe}R_e \leq 0.1$

	CE	CE WITH R_e UNBYPASSED	CC	CB
A_i	$-h_{fe}$	$-h_{fe}$	$1 + h_{fe}$	$\frac{h_{fe}}{1 + h_{fe}}$
R_i	h_{ie}	$h_{ie} + (1 + h_{fe})R_e$	$h_{ie} + (1 + h_{fe})R_L$	$\frac{h_{ie}}{1 + h_{fe}}$
A_v	$\frac{-h_{fe}R_L}{h_{ie}}$	$\frac{-h_{fe}R_L}{R_i}$	$1 - \frac{h_{ie}}{R_i}$	$\frac{h_{fe}R_L}{h_{ie}}$
R_o	∞	∞	$\frac{R_g + h_{ie}}{1 + h_{fe}}$	∞
R_o'	R_L	R_L	$R_o \| R_L$	R_L

TABLE 14-7 COMPARISON OF AMPLIFIER CHARACTERISTICS CALCULATED USING THE APPROXIMATE EQUATIONS OF TABLE 14-6, WITH $R_g = 600$ OHMS, $R_L = 1.5$ K OHMS, $h_{ib} = 20$ OHMS, $h_{rb} = 3 \times 10^{-4}$, $h_{fb} = -0.98$, $h_{ob} = 0.5 \times 10^{-6}$ MHO, OR $h_{ie} = 1$ K OHM $h_{re} = 2 \times 10^{-4}$, $h_{fe} = 49$, $h_{oe} = 25 \times 10^{-6}$ MHO.

QUANTITY	CB	CE	CC
A_i	0.98	-49	50
A_v	78	-78	0.99
R_i ohms	20	1000	76 K
R_o ohms	∞	∞	32
R_o' ohms	1.5 K	1 5 K	31

The values compare well, so that whenever the condition of $h_{oe}R_L \leq 0.1$ is met, the approximate equations will give quite good results, especially when it is realized that the h-parameters are not generally known very accurately anyway. This is because of the variations that exist from one transistor to another, even though of the same type.

14-11. SUMMARY

(1) Any single-stage transistor amplifier, whether CB, CE or CC, may be put in a standard form for small-signal analysis so as to be able to identify V_g, R_g, and R_L as well as the transistor's input and output terminals.

(2) The hybrid equivalent circuit for a transistor consists of a Thevenin equivalent circuit at the input and a Norton equivalent circuit at the output.

(3) The equations relating the input and output voltages and currents in a transistor's small-signal equivalent circuit are given by:

$$V_1 = h_i I_1 + h_r V_2$$
$$I_2 = h_f I_1 + h_o V_2$$

where h_i, h_r, h_f, and h_o are the h-parameters and have a second subscript of b, e, or c according to the amplifier in question.

(4) The h-parameters may be obtained from the transistor's characteristic curves or may be measured experimentally by using changes in current and voltage levels to simulate ac conditions.

(5) Typical h-parameter values for CB, CE, and CC connections are shown in Tables 14-1 and 14-2, and are all dependent upon the collector current.

(6) If the h-parameters are given for one type of amplifier connection but are required for another type, the approximate conversion formulas in Table 14-3 may be used.

(7) Knowing the h-parameters for a transistor, the values for A_i, R_i, A_v, R_o, A_{vg}, and A_{ig} may be obtained using the exact equations in Table 14-4 by using a second subscript appropriate to the amplifier being analyzed. The values of R_g and R_L are also necessary.

(8) Using a transistor's h-parameters, it is possible to dually match R_g and R_L for the maximum available power gain (M.A.G.) from the transistor, using Eqs. 14-14 and 14-15.

(9) A transistor amplifier's characteristics are very dependent upon the values used for R_g and R_L, as shown in Fig. 14-12, but, generally speaking, the characteristics may be summarized as in the following table:

QUANTITY	CB	CE	CC
A_i	Low	High	High
A_v	High	High	Low
R_i	Low	Medium	High
R_o	High	Medium	Low

(10) **Approximate equations for A_i, R_i, A_v, and R_o in terms of the CE h-parameters are shown in Table 14-6 and are valid if $h_{oe}R_L \leq 0.1$.**

QUESTIONS

Q14-1. What variables make up an amplifier circuit that has been converted to standard form for analysis purposes?

Q14-2. What conditions must be met to represent a four-terminal complex circuit by a linear equivalent circuit? How are these conditions met in a transistor?

Q14-3. Why do you think the term "black box" is appropriate in discussing the equivalent circuit of a transistor?

Q14-4. What four units of measurement are found in the Thevenin and Norton hybrid equivalent circuit for a transistor? Why is it appropriate for the output of the equivalent circuit to be described by a Norton equivalent?

Q14-5. Sketch input and output characteristic curves for a transistor in the CB connection and show how you would obtain the four parameters, h_{ib}, h_{fb}, h_{rb} and h_{ob}, as in Eq. 14-3 through 14-6.

Q14-6. Are the h-parameters for a transistor constant? What do they vary with? Which ones show the greatest variation? What other factor determines the values of a transistor's h-parameters?

Q14-7. Which parameters support the fact that a CB has a low input resistance and a CE a high value of input resistance?

Q14-8. Why are h-parameter conversion formulas necessary?

Q14-9. Sketch a PNP CE amplifier and indicate the instantaneous directions of base and collector currents. Now verify that the use of Eq. 14-7, with h_{fe} a positive number, gives the expected directions with reference to Fig. 14-8.

Q14-10. Use Eq. 14-8 to verify that $R_i < h_{ie}$ in a CE, but $R_i > h_{ib}$ in a CB.

Q14-11. Use Eq. 14-9 to verify that phase inversion occurs in a CE, but the phase shift is zero in a CB.

Q14-12. Demonstrate by the use of Eq. 14-11 that the power gain of a CE is positive.

Q14-13. What conditions would make an amplifier's overall voltage gain substantially the same as the transistor's voltage gain? Which amplifier connection is more likely to provide this?

Q14-14. What is meant by dual matching of a transistor in an amplifier?

Q14-15. From an examination of the curves in Fig. 14-12, which amplifier connection shows the greatest isolation between input and output (least variation of R_i with R_L and R_o with R_g)?

PROBLEMS

P14-1. Given the CE small-signal amplifier shown in the diagram and output characteristics for the transistor, do the following:

(a) Draw the dc load line.
(b) Determine the value of R_B to provide a bias current of $60\,\mu A$.
(c) Draw the ac load line for R_o = 1 K ohm. (Assume capacitors have negligible reactance.)
(d) Determine the peak-to-peak output current for a 40 μA p-p input current and obtain the current gain.
(e) Determine the peak-to-peak output voltage for a 40 μA p-p input current. Given that the necessary input voltage from base-to-emitter is 0.025 V p-p, obtain the voltage gain.

(f) Convert the given CE amplifier circuit to standard form so that it is ready for a small-signal analysis, as in Fig. 14-1 (b).

(g) Evaluate h_{fe} in the region of the operating point.

(h) Evaluate h_{oe} in the region of the operating point.

(i) Given that $h_{ie} = 600$ ohms and $h_{re} = 125 \times 10^{-6}$, draw the small-signal equivalent circuit of the CE amplifier showing the numerical values, including R_g and $R_o \| R_L$.

(j) Calculate A_i, R_i, A_v, and G_o using the equations in Table 14-4. Compare the calculated values of A_i and A_v with the graphical values from (d) and (e).

(k) Determine the value of $h_{oe}R_L$. Is it permissible to use the approximate equations of Table 14-6?

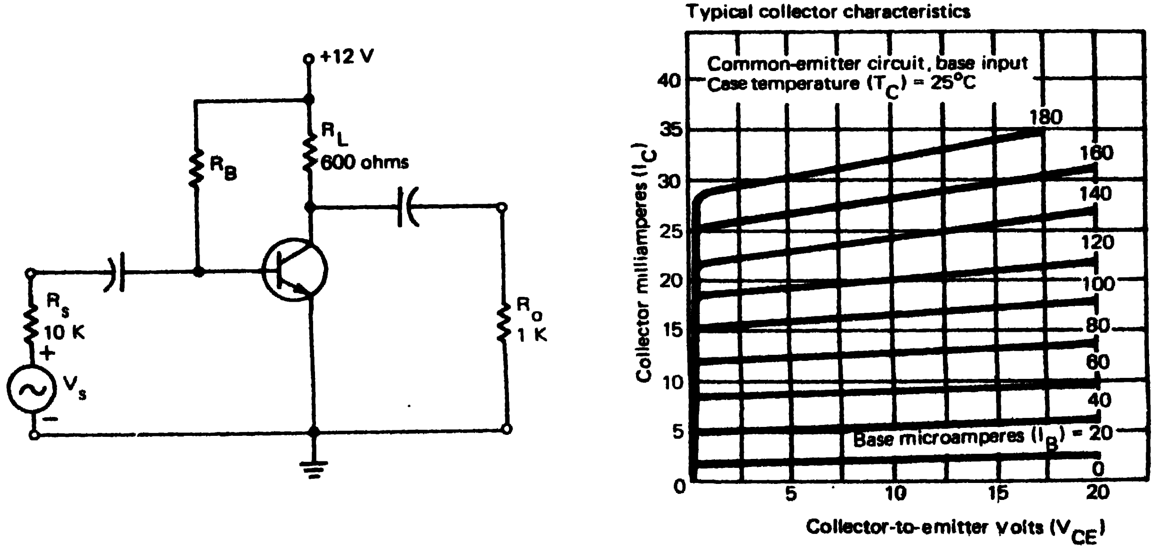
(l) Recalculate A_i, R_i, and A_v, and compare with the values in (j).

(m) Determine A_{vg} and A_{ig} for the overall amplifier.

P14-2. A 2N3391 transistor is used in a CE amplifier at a quiescent collector current of 1 mA. If the load resistance is 5.6 K ohms and the source resistance 600 ohms, calculate A_i, R_i, A_v, A_p, R_o, and R_o'. Use the h-parameters in Table 14-1 and the approximate equations in Table 14-6. (Why may these be used?)

P14-3. A 2N3391 transistor is used in a CE amplifier at a quiescent collector current of 0.1 mA. If the load resistance is 56 K ohms and the source resistance 600 ohms, calculate A_i, R_i, A_v, A_p, R_o, and R_o'. Use the normalized curves in Fig. 14-7 and the h-parameters in Table 14-1 to obtain the appropriate h-parameters. Use the exact equations in Table 14-4. (Why do these have to be used?)

P14-4. Calculate the values of source and load resistance to provide dual matching for maximum power gain in P14-3. Determine A_v, A_i, and A_p for this condition.


+12 V
R_B
R_L
600 ohms
R_s
10 K
+
V_s
−
R_o
1 K
Typical collector characteristics
Common-emitter circuit, base input
Case temperature (T_C) = 25°C
Collector milliamperes (I_C)
0
5
10
15
20
25
30
35
40
180
160
140
120
100
80
60
40
Base microamperes (I_B) = 20
0
5
10
15
20
Collector-to-emitter volts (V_{CE})

CE amplifier and collector characteristics for P14-1. *(Courtesy RCA Corp.)*

P14-5. A transistor with h-parameters as given in Table 14-2 is to be used in a CC amplifier to provide an input resistance of 150 K ohms.

(a) Use the approximate equations in Table 14-6 to determine the necessary value of R_L.

(b) Calculate A_i, A_v, and R_o if the amplifier is driven from a 25-K ohm source.

(c) Justify being able to use the approximate equations.

P14-6. A transistor with h-parameters as given in Table 14-2 is used in a CE amplifier that has a load resistance of 5.6 K ohms and an unbypassed 100-ohm resistor in the emitter lead.

(a) Use the approximate equations in Table 14-6 to determine A_i, R_i, A_v, and R_o'.

(b) Recalculate R_i and A_v if R_e is bypassed.

P14-7. A 2N718A transistor is used in a CB amplifier with R_L = 150 K ohms and R_s = 10 ohms. The h-parameters given are h_{ib} = 40 ohms, $h_{rb} = 3 \times 10^{-4}$, h_{ob} = 1 μmho, and h_{fe} = 60. Calculate A_i, R_i, A_v, A_p, R_o, R_o', A_{vg}, and A_{ig}.

P14-8. Determine the values of R_L and R_s (R_g) to provide dual matching in P14-7 and obtain A_v, A_i, and A_p for this condition.

P14-9. Use the conversion formulas in Table 14-3 to verify the CE h-parameters in the caption of Table 14-7, which were obtained from the given CB h-parameters.

P14-10. The 2N718A transistor in P14-7 is used in a CE amplifier with R_L = 10 K ohms. and driven from a source of 1 K ohm. Calculate: A_i, R_i, A_v, A_p, R_o, and R_o'.

P14-11. A 2N1302 has the following h-parameters given: h_{ib} = 28 ohms, $h_{rb} = 5 \times 10^{-4}$, h_{ob} = 0.34 μmhos, and h_{fe} = 105. Calculate A_i, R_i, and A_v for a common-collector amplifier if R_L = 1.2 K ohms.

P14-12. A 2N3391 is operated at a collector current of 20 mA in a CB circuit with R_L = 3.3 K ohms.

(a) Use the data in Table 14-1 and the curves in Fig. 14-7 to obtain the CE h-parameters.

(b) Use the conversion formulas in Table 14-3 to obtain the CB h-parameters.

(c) Calculate A_i, R_i, and A_v using the exact equations in Table 14-4 and CB h-parameters.

(d) Calculate A_i, R_i and A_v using the approximate equations in Table 14-6 and the CE h-parameters. What minimum value of R_L would be necessary to make it essential to use the exact equations in Table 14-4?

15

LOW-FREQUENCY RESPONSE OF THE TRANSISTOR AMPLIFIER

15.0 INTRODUCTION

In the previous chapter, we developed equations for the current and voltage gains for small-signal amplifiers operating at low to mid-range frequencies. We assumed that the reactance of coupling and bypass capacitors was negligible at the frequency of operation. We now examine the factors affecting the low-frequency response of a single stage common-emitter amplifier and the modifications necessary when considering a number of capacitor-coupled CE stages in cascade. Finally, the transformer-coupled amplifier is analyzed for its low-frequency response, and the problem of sag in square-wave amplifiers is considered.

15-1. EFFECT OF AN EMITTER BYPASS CAPACITOR ON LOW-FREQUENCY RESPONSE

We saw in Chapter 12 how the addition of a resistor in the emitter lead of a CE amplifier contributed to biasing and thermal stability in the amplifier. We also indicated that this resistor should be adequately bypassed to avoid a degenerative effect due to negative feedback. Let us now determine the effect of the bypass capacitor, C_e, shown in Fig. 15-1(a), on the low-frequency response.

It is required to find an expression for $A_{v(LF)} = \frac{V_o}{V_s}$, assuming that C_c is sufficiently large so as to not have any effect on the response. The effect of C_c is considered in the next section. Also, the parallel combination of R_1 and R_2 is assumed to be much larger than R_i. If it is not, the actual voltage gain will be less than that predicted by the following, but its effect is taken into account in the equations in Table 15-1.

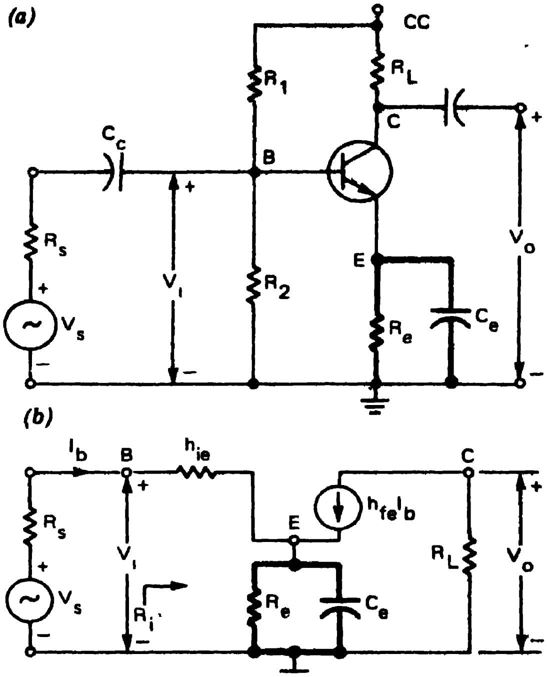


Fig. 15-1. (a) CE amplifier to be analyzed for the effect of C_e on the low-frequency response.
(b) Approximate small-signal hybrid equivalent circuit at low frequencies for the CE amplifier.

With reference to Fig. 15-1(b), $V_o = -h_{fe}I_bR_L$. But

$$I_b = \frac{V_s}{R_s + R_i},$$

where, from Table 14-6 for a CE with an unbypassed R_e, $R_i = h_{ie} + (1 + h_{fe})Z_e$,

where $Z_e = R_e \| C_e = \dfrac{R_e}{1 + j\omega C_e R_e}$.

Thus, $$V_o = -h_{fe}R_L \times \frac{V_s}{R_s + h_{ie} + \dfrac{(1 + h_{fe})R_e}{1 + j\omega C_e R_e}}.$$

and $$A_{v(LF)} = \frac{V_o}{V_s} = \frac{-h_{fe}R_L}{R_s + h_{ie} + \dfrac{(1 + h_{fe})R_e}{1 + j\omega C_e R_e}} \qquad (15\text{-}1)$$

Note that when ω is large (in the mid-frequency range) the voltage gain,

$$A_{v(MF)} = \frac{-h_{fe}R_L}{R_s + h_{ie}},$$

since the third term in the denominator of Eq. 15-1 tends to zero. This substantitates the findings in Chapter 14.

The ratio of the gain at low frequencies to the gain at mid frequencies,

$\dfrac{A_{v(LF)}}{A_{v(MF)}}$, may be put in the following form:

$$\frac{A_{v(LF)}}{A_{v(MF)}} = \frac{1}{1 + \dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}}} \times \frac{1 + j\dfrac{f}{f_o}}{1 + j\dfrac{f}{f_p}}, \qquad (15\text{-}2)$$

where $f_o = \dfrac{1}{2\pi C_e R_e}$, and $f_p = \dfrac{1 + \dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}}}{2\pi C_e R_e}$.

If $\dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}} >> 1$, then $f_p \approx \dfrac{\dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}}}{2\pi C_e R_e}$.

This means that $f_p >> f_o$, so that if we consider $\dfrac{A_{v(LF)}}{A_{v(MF)}}$ at $f = f_p$ we obtain

$$\frac{A_{v(LF)}}{A_{v(MF)}} = \frac{R_s + h_{ie}}{(1 + h_{fe})R_e} \times \frac{1 + j\dfrac{f_p}{f_o}}{1 + j1}$$

$$\approx \frac{R_s + h_{ie}}{(1 + h_{fe})R_e} \times \frac{j\dfrac{f_p}{f_o}}{1 + j1}.$$

The magnitude of this ratio is

$$\left|\frac{A_{v(LF)}}{A_{v(MF)}}\right| \approx \frac{R_s + h_{ie}}{(1 + h_{fe})R_e} \times \frac{\dfrac{f_p}{f_o}}{\sqrt{2}}.$$

But $\dfrac{f_p}{f_o} \approx \dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}}$.

$\therefore \left|\dfrac{A_{v(LF)}}{A_{v(MF)}}\right| = \dfrac{1}{\sqrt{2}}$, which means the power gain at this low frequency has dropped 3 dB from the gain at the mid frequency.

Thus the lower 3-dB frequency is

$$f_1 \approx f_p = \frac{(1 + h_{fe})R_e}{(R_s + h_{ie})2\pi C_e R_e}$$

$$= \frac{1 + h_{fe}}{2\pi C_e(R_s + h_{ie})} \qquad (15\text{-}3)$$

when $f_p >> f_o$. If this condition is not met, $f_1 \neq f_p$, and there may not be a 3-dB point. This will occur if

$$\frac{(1 + h_{fe})R_e}{R_s + h_{ie}} < 0.414. \qquad (15\text{-}4)$$

Note that the expression for f_1 does not include R_e, so that the choice of a bypass capacitor for a given lower 3-dB frequency is dependent only upon the transistor and the source resistance.

Example 15-1. Calculate the size of bypass capacitor, C_e, to provide a low-frequency 3-dB point of 100 Hz when $R_e = 1$ K ohm, $h_{fe} = 50$, $h_{ie} = 1$ K ohm, and $R_s = 600$ ohms.

Solution. Using Eq. 15-3,

$$C_e = \frac{1 + h_{fe}}{2\pi(R_s + h_{ie})f_1}$$

$$\therefore C_e = \frac{1 + 50}{2\pi(600 + 1000)100}\ \text{F} = 50\ \mu\text{F}.$$

To justify the use of Eq. 15-3, determine f_o and f_p:

$$f_o = \frac{1}{2\pi C_e R_e} = \frac{1}{2\pi 50 \times 10^{-6} \times 1 \times 10^3} = 3.2\ \text{Hz}.$$

$$f_p = \frac{1 + \dfrac{(1 + h_{fe})R_e}{R_s + h_{ie}}}{2\pi C_e R_e} = \frac{1 + \dfrac{(1 + 50)1000}{600 + 1000}}{2\pi 50 \times 10^{-6} \times 1000}$$

$$= 105\ \text{Hz}.$$

Clearly, $f_p >> f_o$, and so $f_1 \approx f_p$, as given by Eq. 15-3.

It should be noted that the use of electrolytic capacitors to provide the high capacitance required for C_e may cause a reduction in the lower 3-dB frequency below 100 Hz. This is due to the series electrolyte loss resistance of the capacitor, which may be anywhere from 10 to 20 ohms. This resistance also reduces the mid-frequency gain, so that if R_C represents the series resistance of C_e, $A_{v(MF)}$ is given by

$$A_{v(MF)} \approx \frac{-h_{fe}R_L}{R_s + h_{ie} + (1 + h_{fe})R_C}. \qquad (15\text{-}5)$$

The percent reduction in gain due to R_C is approximately equal to the percent reduction in the low-frequency 3-dB point, f_1, as shown in Example 15-2.

Example 15-2. Using the amplifier and conditions stated in Example 15-1, calculate the mid-frequency gain if $R_L = 2$ K ohms under the following conditions:

(a) Using a 50-μF capacitor with no series resistance.

(b) Using a 50-μF capacitor which has a 15-ohm series resistance.

Also determine the new value of f_1 with the capacitor as in (b).

Solution

(a) With $R_C = 0$, $A_{v(MF)} = \dfrac{-h_{fe}R_L}{R_s + h_{ie}}$

$$= \frac{-50 \times 2{,}000}{600 + 1{,}000} = -63.$$

(b) With $R_C = 15$ ohms,

$$A_{v(MF)} = \frac{-h_{fe}R_L}{R_s + h_{ie} + (1 + h_{fe})R_C}$$

$$= \frac{-50 \times 2{,}000}{600 + 1000 + (1 + 50)15}$$

$$= \frac{-100{,}000}{1600 + 765} = -42.$$

With $R_C = 0$, $A_{v(MF)} = -63$, $f_1 = 100$ Hz.

With $R_C = 15$ ohms, $A_{v(MF)} = -42$, $f_1 = ?$

$$\frac{f_1}{100} = \frac{-42}{-63} \therefore f_1 = 100 \times \frac{42}{63} = 67\ \text{Hz}.$$

Although the reduction in value of f_1 from 100 Hz to 67 Hz is not disagreeable, the significant drop in $A_{v(MF)}$, from -63 to -42 is objectionable. Accordingly, there are available special transistor bypass capacitors of high capacitance, low working voltage, and a low series resistance of approximately 1 ohm, which may be used for C_e.

If the effect of the biasing resistors, R_1 and R_2, is included, the mid-frequency voltage gain is given by $A_{v(MF)}$

$$\approx \frac{-h_{fe}R_L}{R_s + h_{ie} + \frac{h_{ie}R_s}{R}}, \text{ where } R = R_1 \parallel R_2. \qquad (15\text{-}6)$$

The combined effect of R_C and R on the mid-frequency gain, along with the corresponding equations for f_1 under these conditions, is indicated in Table 15-1.

Note: The equations developed above for mid-frequency voltage gain and f_1 were for $\frac{V_o}{V_s}$. The equations for $\frac{V_o}{V_i}$ may be obtained from these equations by setting $R_s = 0$ and are listed in Table 15-1. The corresponding equations for f_1 where $\frac{V_o}{V_i}$ drops by 3 dB from the mid-frequency value may also be obtained by setting $R_s = 0$ in the equations for f_1 in Table 15-1.

TABLE 15-1 SUMMARY OF THE APPROXIMATE EQUATIONS FOR MID-FREQUENCY VOLTAGE GAIN AND LOWER 3-dB FREQUENCY, f_1, FOR VARIOUS CIRCUIT CONDITIONS

CONDITION	$A_{v(MF)} = \frac{V_o}{V_s}$	$A_{v(MF)} = \frac{V_o}{V_i}$	f_1, Hz, for $\frac{V_o}{V_s}$
$R_1 \parallel R_2 \gg R_i$ $R_C = 0$	$\frac{-h_{fe}R_L}{R_s + R_{i1}}$	$\frac{-h_{fe}R_L}{R_{i1}}$	$\frac{1 + h_{fe}}{2\pi C_e(R_s + R_{i1})}$
$R_1 \parallel R_2 \gg R_i$ $R_C \neq 0$	$\frac{-h_{fe}R_L}{R_s + R_{i2}}$	$\frac{-h_{fe}R_L}{R_{i2}}$	$\frac{1 + h_{fe}}{2\pi C_e(R_s + R_{i2})}$
$R_1 \parallel R_2 \not\gg R_i$ $R_C = 0$	$\frac{-h_{fe}R_L}{R_s + R_{i3}}$	$\frac{-h_{fe}R_L}{R_{i1}}$	$\frac{1 + h_{fe}}{2\pi C_e(R_s + R_{i3})}$
$R_1 \parallel R_2 \not\gg R_i$ $R_C \neq 0$	$\frac{-h_{fe}R_L}{R_s + R_{i4}}$	$\frac{-h_{fe}R_L}{R_{i2}}$	$\frac{1 + h_{fe}}{2\pi C_e(R_s + R_{i4})}$

(See the following page for resistance designations.)

$$R_{i1} = h_{ie}$$

$$R_{i2} = h_{ie} + (1 + h_{fe})R_C$$

$$R_{i3} = h_{ie} + \frac{h_{ie}R_s}{R_1 \| R_2}$$

$$R_{i4} = h_{ie} + \frac{h_{ie}R_s}{R_1 \| R_2} + (1 + h_{fe})R_C$$

R_C = series resistance of the electrolytic capacitor, C_e. Equations for f_1 are valid provided $\frac{(1 + h_{fe})R_e}{R_s + h_{ie}} >> 1$.

15-2. EFFECT OF A COUPLING CAPACITOR ON LOW-FREQUENCY RESPONSE

We now assume that C_e in Fig. 15-2(a) is large enough to cause no reduction in low-frequency gain, so that the effect of the coupling capacitor, C_c, may be examined. A low-frequency model for the CE amplifier is shown in Fig. 15-2(b).

The mid-frequency gain, $A_{v(MF)} = \frac{V_o}{V_i}$, is given by the approximate equations of Table 15-1 since the reactance of C_c is negligible. At low frequencies, however, the voltage drop across C_c will reduce V_i, with a corresponding drop in V_o. The frequency at which the gain drops by a factor of $\frac{1}{\sqrt{2}}$, the lower 3-dB frequency, f_1, is given by

$$f_1 = \frac{1}{2\pi(R_s + R_i')C_c} \tag{15-7}$$

where $R_i' = R_1 \| R_2 \| R_i$, and $R_i \approx h_{ie}$ for an ideal capacitor C_e, or $R_i \approx h_{ie} + (1 + h_{fe})R_C$ if the capacitor's series resistance is taken into account.

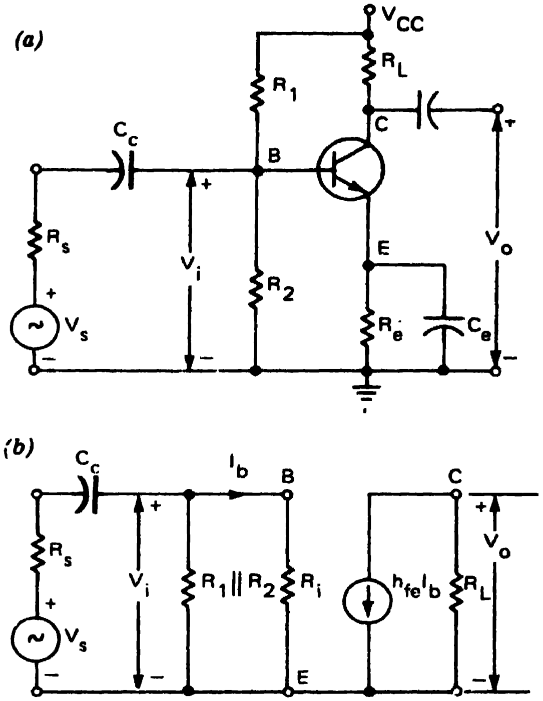


Fig. 15-2. (a) CE amplifier to be analyzed for the effect of C_c on the low-frequency response.
(b) Low frequency model for a CE amplifier showing the coupling capacitor, C_c.

Example 15-3. Calculate the size of coupling capacitor C_c in Fig. 15-2(a) to provide a low-frequency 3-dB point at 100 Hz if R_s = 600 ohms, h_{ie} = 1 K ohm, h_{fe} = 50, R_1 = 5.2 K ohms, R_2 = 1.24 K ohms. Assume:

(a) An ideal bypass capacitor, R_C = 0.
(b) A practical bypass capacitor, R_C = 15 ohms.

Solution

(a) From Eq. 15-7,

$$C_c = \frac{1}{2\pi f_1(R_s + R_i')}.$$

$$R_i' = R_1 \| R_2 \| R_i = R_1 \| R_2 \| h_{ie}$$
$$= 5.2\text{ K} \| 1.24\text{ K} \| 1\text{ K} = 500 \text{ ohms.}$$

$$\therefore C_c = \frac{1}{2\pi 100(600 + 500)} \text{ F} = 1.45\ \mu\text{F}.$$

(b) $R_i' = R_1 \| R_2 \| (h_{ie} + (1 + h_{fe})R_C)$

$= 5.2 \text{ K} \| 1.24 \text{ K} \| 1.77 \text{ K}$

$= 640$ ohms.

$$\therefore C_c = \frac{1}{2\pi 100(600 + 640)} \text{ F} = 1.26\ \mu\text{F}.$$

If the required value for f_1 is 10 Hz instead of 100 Hz, the above capacitors would be 14.5 and 12.8 μF, respectively. These relatively large values are a result of the transistor's low input resistance (and biasing resistors) and are much larger than the coupling capacitors required for tube or FET amplifiers. A method of increasing the input resistance will be covered in Chapter 17.

If the calculation of f_1 due to C_c is different from the value of f_1 due to C_e by a factor of 5 or more, then the overall value for f_1 is approximately the higher of the two. The overall value of f_1 when they are equal will be covered in a later section.

It should be evident from the previous two sections that large values of capacitors are required in transistor amplifiers for coupling and bypass purposes for good low frequency response. Since these capacitors are bulky, audio amplifiers utilizing integrated circuits often have large capacitors connected externally to the IC itself to provide the requisite low-frequency response.

15-3. CASCADING OF CE STAGES, MID-FREQUENCY GAINS

To build up a signal's input voltage and current to a usable level, it is often necessary to cascade stages. That is, use one amplifier to feed another and so on. Since the CE has the highest power gain, it is the most popular for cascading, but a mixture of the three common types may be used to advantage for matching purposes. For example, a common-collector with its low output impedance may be used as the last stage, so that it may drive low resistance loads. And the common-collector with its high input impedance may also be a good circuit for the first stage. In any event, the overall voltage gain of an amplifier with many stages will be the *product* of the stage gains. If the stage gains are computed in decibels, the overall gain is the *sum* of the stage gains, and the total phase shift the sum of the phase shifts introduced by each stage.

Consider obtaining expressions for the mid-frequency voltage and current gains of the three stages shown in Fig. 15-3, using the approximate equations for a CE from Table 14-6. Assume C_e and C_c have negligible reactance and resistance.

Since the input resistance of the third stage determines the voltage gain of the second we must start at the last stage and work backwards from V_o to V_1.

Third Stage

$$A_{i3} = \frac{I_{c3}}{I_{b3}} = -h_{fe} = -50.$$

$$R_{i3} = h_{ie} = 1 \text{ K ohm}.$$

$$A_{v3} = \frac{V_o}{V_3} = \frac{A_{i3}R_{L3}}{R_{i3}} = \frac{-50 \times 2 \times 10^3}{1 \times 10^3} = -100.$$

Second Stage

$$A_{i2} = \frac{I_{c2}}{I_{b2}} = -h_{fe} = -50.$$

$$R_{i2} = h_{ie} = 1 \text{ K ohm}.$$

The effective load for Q2 is the parallel combination of R_{L2}, $R_1 \| R_2$ (referred to as R), and R_{i3}. That is, R_{L2}' is given by

$$R_{L2}' = R_{L2} \| R \| R_{i3} = 2 \text{ K} \| 1 \text{ K} \| 1 \text{ K} = 400 \text{ ohms}.$$

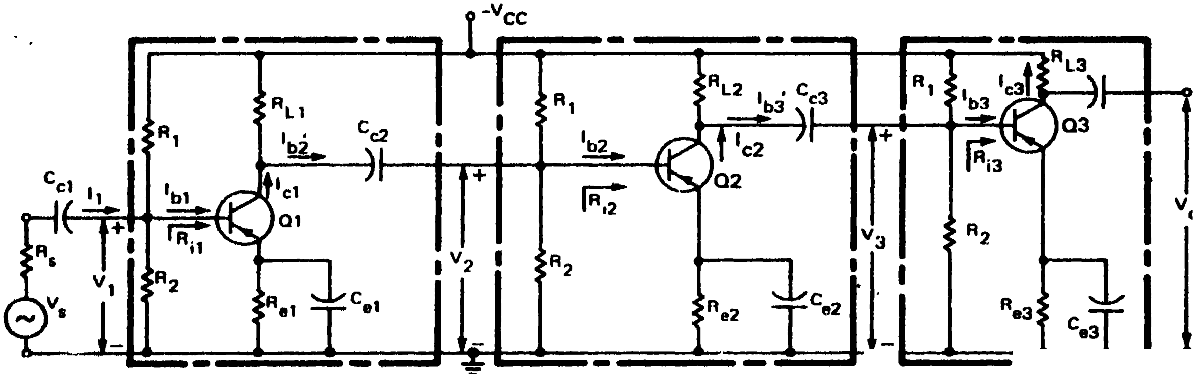


Fig. 15-3. Three CE cascaded stages. $R_{L1} = R_{L2} = R_{L3} = 2$ *K ohms,* $h_{fe} = 50$, $h_{ie} = 1$ *K ohm,* $R_1 = 5.2$ *K ohms,* $R_2 = 1.24$ *K ohms,* $R = R_1 \| R_2 = 1$ *K ohm,* $R_s = 600$ *ohms,* $C_{c1} = 2\ \mu F$, $C_{c2} = C_{c3} = 1\ \mu F$, $C_{e1} = 50\ \mu F$, $C_{e2} = C_{e3} = 25\ \mu F$.

$$A_{v2} = \frac{V_3}{V_2} = \frac{A_{i2}R_{L2}'}{R_{i2}} = \frac{-50 \times 400}{1000} = -20.$$

Note the damaging effect on voltage gain of the low input impedance of the next stage and its biasing resistors.

First Stage

$$A_{i1} = \frac{I_{c1}}{I_{b1}} = -h_{fe} = -50, \text{ and } R_{i1} = h_{ie} = 1 \text{ K ohm.}$$

$$R_{L1}' = R_{L2}' = 400 \text{ ohms.}$$

$$A_{v1} = \frac{V_2}{V_1} = \frac{A_{i1}R_{L1}'}{R_{i1}} = \frac{-50 \times 400}{1000} = -20.$$

Overall Gains

$$A_v = \frac{V_o}{V_1} = \frac{V_o}{V_3} \times \frac{V_3}{V_2} \times \frac{V_2}{V_1}$$

$$= -100 \times (-20) \times (-20) = -40{,}000.$$

Voltage gain in dB = 20 log 100 + 20 log 20 + 20 log 20 = 40 + 26 + 26 = 92 dB, or, 20 log 40,000 = 20 × 4.6 = 92 dB.

Note: The above expression for voltage gain in dB assumes equal input and output resistances, which is not true in this case. However, common usage is to call 20 log A_v the dB voltage gain even if it is not a true dB power gain = 10 log A_p.

If we require $A_{vs} = \frac{V_o}{V_s}$, we must take into account R_s, R, and R_{i1}. Thus,

$$V_1 = V_s \times \frac{R\|h_{ie}}{R_s + R\|h_{ie}} = V_s \times \frac{5}{11}.$$

$$A_{vs} = A_v \times \frac{V_1}{V_s} = -40{,}000 \times \frac{5}{11} = -18{,}200.$$

Although the overall voltage gain was obtained from the product of the stage gains, the overall current gain is not obtainable in the same way. This is because the output current of one stage is not the input current for the next. There is a division of current as shown by the partial circuit in Fig. 15-4 for the second stage.

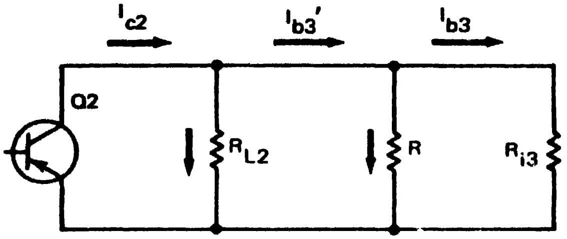


Fig. 15-4. Showing how the current divides in the output circuit of a CE stage.

From Fig. 15-4, $\frac{I_{b3}}{I_{c2}} = \frac{R_{L2}\|R}{R_{L2}\|R + R_{i3}} = \frac{2}{5}$.

$$\therefore A_i' = \frac{I_{c3}}{I_{b1}} = \frac{I_{c3}}{I_{b3}} \times \frac{I_{b3}}{I_{c2}} \times \frac{I_{c2}}{I_{b2}} \times \frac{I_{b2}}{I_{c1}} \times \frac{I_{c1}}{I_{b1}}$$

$$= A_{i3} \times \frac{2}{5} \times A_{i2} \times \frac{2}{5} \times A_{i1}$$

$$= (-50) \times \frac{2}{5} \times (-50) \times \frac{2}{5} \times (-50)$$

$$= -20{,}000.$$

$$A_i = \frac{I_{c3}}{I_1} = A_i' \times \frac{I_{b1}}{I_1} = A_i' \times \frac{R}{R + R_{i1}}$$

$$= -20{,}000 \times \frac{1}{2} = -10{,}000.$$

$$A_p = A_v \times A_i = (-40{,}000) \times (-10{,}000)$$

$$= 4 \times 10^8.$$

The power decibel gain = 10 log 4 × 10^8 = 10 × 8.6 = 86 dB. Note that the voltage gain may also be obtained from

$$A_v = \frac{A_i' R_{L3}}{R_{i1}} = \frac{-20{,}000 \times 2 \times 10^3}{1 \times 10^3} = -40{,}000,$$

and

$$A_{vs} = \frac{A_i \times R_{L3}}{R_s + R\|R_{i1}}$$

$$= \frac{-10{,}000 \times 2 \times 10^3}{600 + 500} = -18{,}200.$$

Thus, a general expression for overall voltage gain from input to output is the product of the current gain and the output load resistance, divided by the input resistance at the point where the current gain was computed.

The above figures for A_v and A_i will be within 10 percent of the actual values if $h_{oe}R_L \leq 0.1$, as required by Table 14-6.

15-4. LOW-FREQUENCY RESPONSE OF CASCADED STAGES

To determine the overall lower 3-dB frequency for V_o/V_s, we obtain f_1 for the coupling and bypass capacitors separately and then combine the results.

Coupling Capacitors

Let us now determine the low-frequency response, f_1, for each stage due to the coupling capacitors with values as shown in Fig. 15-3.

Using Eq. 15-7 from Section 15-2, we have:

First Stage

$$f_1 = \frac{1}{2\pi(R_s + R_i')C_{c1}}$$

$$= \frac{1}{2\pi(600 + 500)2 \times 10^{-6}} = 72 \text{ Hz}.$$

Second Stage. The effective source resistance R_s for the input to the second stage is the output resistance R_o' of the first stage, which, from the approximate equations of Table 14-6, equals $R_{L1} = 2$ K ohms.

$$\therefore f_1 = \frac{1}{2\pi(R_o' + R_i')C_{c2}}$$

$$= \frac{1}{2\pi(2000 + 500)1 \times 10^{-6}} = 64 \text{ Hz}.$$

Third Stage. The effect of C_{c3} will be the same as that for C_{c2}, giving an f_1 for the third stage of 64 Hz.

Overall Response. It may be shown that the overall low-frequency response for n cascaded stages each having the same value for f_1 is given by the frequency $f_{1(n)}$ where the overall voltage gain falls to $1/\sqrt{2}$ (3 dB) of its mid-frequency value. That is,

$$f_{1(n)} = \frac{f_1}{\sqrt{2^{1/n} - 1}}. \qquad (15\text{-}8)$$

Let us assume that $f_1 = 64$ Hz for all three stages in our example. Then the overall 3-dB frequency is

$$f_1 = \frac{64}{\sqrt{2^{\frac{1}{3}} - 1}} = \frac{64}{\sqrt{\sqrt[3]{2} - 1}} = \frac{64}{\sqrt{1.26 - 1}}$$

$$= \frac{64}{\sqrt{0.26}} = \frac{64}{0.51} = 125 \text{ Hz.}$$

Emitter Bypass Capacitors

The value of f_1 due to the emitter bypass capacitors may be obtained using the equation for f_1 from Table 15-1.

That is,

$$f_1 = \frac{1 + h_{fe}}{2\pi C_e (R_s + R_{i3})}$$

where

$$R_{i3} = h_{ie} + \frac{h_{ie} R_s}{R_1 \| R_2}.$$

First Stage

$$R_s + R_{i3} = 0.6\text{ K} + 1\text{ K} + \frac{1\text{ K} \times 0.6\text{ K}}{1\text{ K}} = 2.2\text{ K ohms.}$$

$$f_1 = \frac{1 + 50}{2\pi 50 \times 10^{-6} \times 2.2 \times 10^3} = 74 \text{ Hz.}$$

Second and Third Stages

$$R_s + R_{i3} = 2\text{ K} + 1\text{ K} + \frac{1\text{ K} \times 2\text{ K}}{1\text{ K}} = 5\text{ K ohms.}$$

$$f_1 = \frac{1 + 50}{2\pi 25 \times 10^{-6} \times 5 \times 10^3} = 65 \text{ Hz.}$$

Overall Response

Assuming $f_1 = 64$ Hz for each stage, we obtain an overall value of $f_1 = 125$ Hz, the same as for the coupling capacitors alone.

Combined Effect of Coupling and Emitter-Bypass Capacitors

The combined effects of C_c and C_e may be obtained by using Eq. 15-8 with $f_1 = 125$ Hz and $n = 2$.

$$f_1 = \frac{125}{\sqrt{2^{\frac{1}{2}} - 1}} = \frac{125}{0.64} = 194 \text{ Hz.}$$

Alternatively, the combined effects of C_c and C_e on each stage provide an $f_1 = 100$ Hz for each stage, and then the use of Eq. 15-8 with $f_1 = 100$ Hz and $n = 3$ gives f_1 overall = 194 Hz.

15-5. AMPLIFIER LOW-FREQUENCY RESPONSE TO A SQUARE WAVE

The significance of the lower 3-dB frequency, f_1, may be seen by considering the application of a low-frequency square wave to the circuit of Fig. 15-5(a). This is a representation of a single CE amplifier stage at low frequencies and should be compared with Fig. 15-2. It has a lower 3-dB frequency given by

$$f_1 = \frac{1}{2\pi R_1 C_1}, \qquad (15\text{-}9)$$

which is of the same form as Eq. 15-7.

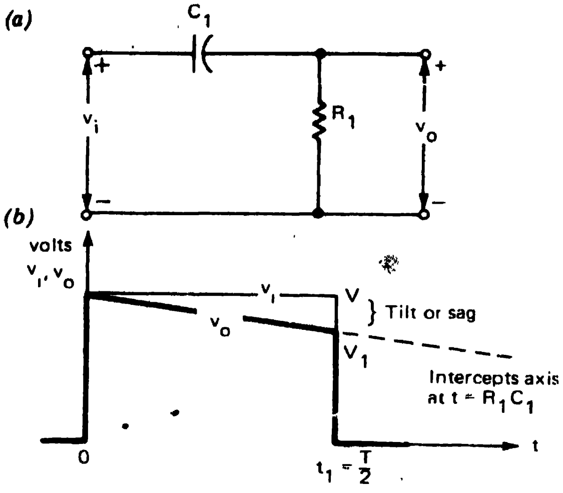


Fig. 15-5. (a) Simplified low-frequency model of a CE amplifier for tilt calculations. (b) Response of the circuit to a step-voltage input.

When a step voltage of amplitude V is applied for v_i, the charging of C_1 will cause the output, v_o, to decrease from V to V_1 after half the period of the square wave has elapsed, given by $t_1 = \frac{T}{2}$.

The difference in voltage, $V - V_1$, is the sag or tilt and is expressed as a percentage of the step input V as follows:

$$\text{Sag} = \frac{V - V_1}{V} \times 100\%. \qquad (15\text{-}10)$$

The expression for v_o is given as a function of time by

$$v_o = Ve^{\frac{-t}{R_1C_1}}$$

so that at $t = t_1$, $v_o = V_1$.

However, recall that the *initial* rate of decrease of v_o is at a rate that would reduce v_o to zero in one time constant, R_1C_1. So, for small values of t_1 compared with R_1C_1, v_o decreases approximately linearly for the half period t_1. Thus, by similar triangles,

$$\frac{V - V_1}{t_1} = \frac{V}{R_1C_1}. \qquad (15\text{-}11)$$

$$\text{From which } \frac{V - V_1}{V} = \frac{t_1}{R_1C_1}. \qquad (15\text{-}12)$$

Since $t_1 = \frac{T}{2} = \frac{1}{2f}$, where f is the frequency of the square wave, and $R_1C_1 = \frac{1}{2\pi f_1}$ from Eq. 15-9, we obtain

$$\text{Sag} = \frac{V - V_1}{V} = \frac{\pi f_1}{f} \times 100\%. \qquad (15\text{-}13)$$

Rearranging for f_1, we obtain $f_1 = \frac{f}{\pi} \times \text{Sag}$.

(15-14;

This indicates that, if an amplifier is to reproduce a low-frequency square-wave with a low degree of sag (high-quality reproduction), a very low value of f_1 is required. For example, to reproduce a 100 Hz square wave with not more than 10% sag requires a lower 3-dB frequency, f_1, of $100/\pi \times 0.1 = 3.2$ Hz.

To get some idea of the size of capacitors needed to provide this low value of f_1, assume that C_c has been chosen large enough to make f_1 dependent upon C_e only. In Example 15-1, a 50-μF capacitor was required to provide an $f_1 = 100$ Hz. Using Eq. 15-3, the necessary value of C_e for $f_1 = 3.2$ Hz, using the same values for h_{ie}, h_{fe}, and R_s is

$$C_e = \frac{1 + h_{fe}}{2\pi(R_s + h_{ie})f_1} = \frac{1 + 50}{2\pi(600 + 1000)3.2}$$

$$= 1600\ \mu\text{F}.$$

This is an impractically large value. The only recourse is to leave the emitter unbypassed and suffer a loss in gain if low-frequency square waves are to be reproduced with little sag. If the loss of gain cannot be tolerated, the emitter resistor would have to be omitted, with a loss of its beneficial characteristics. The mid-frequency gain with R_e unbypassed is given by the equations in Table 15-1 with $R_C = R_e$.

15-6. TRANSFORMER-COUPLED TRANSISTOR AMPLIFIERS

For low-frequency audio amplifiers, the coupling of one stage to another may be made more effective by the use of transformers to match the output impedance of one stage to the input impedance of the next. This is shown in Fig. 15-6.

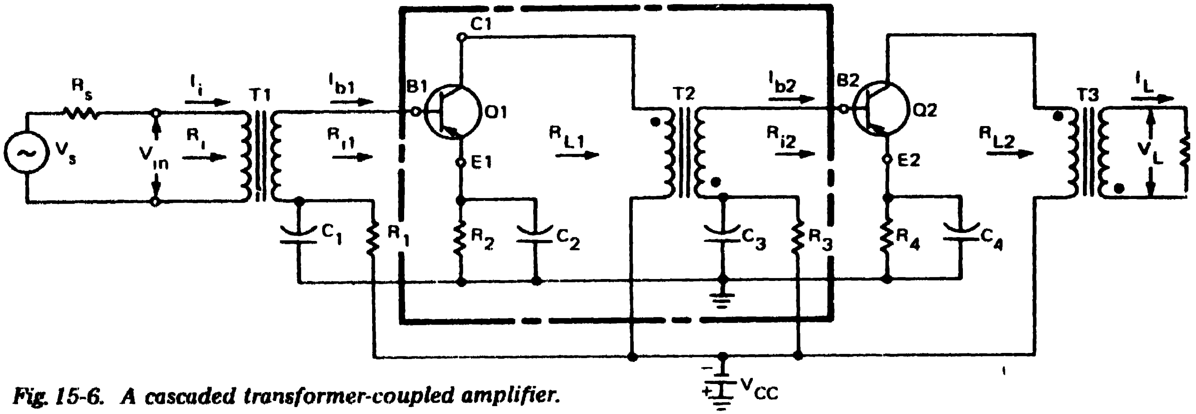


Fig. 15-6. A cascaded transformer-coupled amplifier.

C_1, C_2, C_3, and C_4 are signal bypass capacitors to apply the signal between base and emitter. R_1 and R_3 are resistors used to provide correct biasing, while R_2 and R_4 are emitter stabilizing resistors. Given R_L, the circuit would be analyzed starting at the right, obtaining R_{L2} by taking into account the turns ratio of the *output* transformer, T3.

This would be the load for transistor Q2 and should ideally be equal to the output resistance of Q2 for maximum power transfer. Next, the input resistance to Q2, equal to R_{i2} (h_{ie2}) would permit the determination of the load reflected across the *interstage* transformer T2 if the turns ratio of T2 were known. This would be the effective load R_{L1} for Q1. Finally, the input resistance R_{i1} to Q1, (equal to h_{ie1}), would allow the selection of the *driver* transformer, T1, based on a knowledge of the source impedance, R_s. In general, transformers are not selected for their turns ratio, but for the impedances they are to match and their power handling capability. However, these impedances would permit the determination of the turns ratio for calculating current and voltage gains.

Mid-Frequency Stage Gain, A_i

Let us consider the selection of transformer T2 for the stage enclosed in Fig. 15-6 and the current gain of the stage, $A_i = \frac{I_{b2}}{I_{b1}}$. Assuming that all capacitor reactances and resistances are negligible, we can use the approximate equivalent circuit in Fig. 15-7(a). Recalling that a transformer of turns ratio $\frac{N_1}{N_2}$ reflects a load impedance on the secondary, h_{ie2}, into the primary equal to a value of $\left(\frac{N_1}{N_2}\right)^2 h_{ie2}$, we can use the simplified equivalent circuit of Fig. 15-7(b). Thus, Q1 has an effective load equal to $\left(\frac{N_1}{N_2}\right)^2 h_{ie2}$, and, since its output

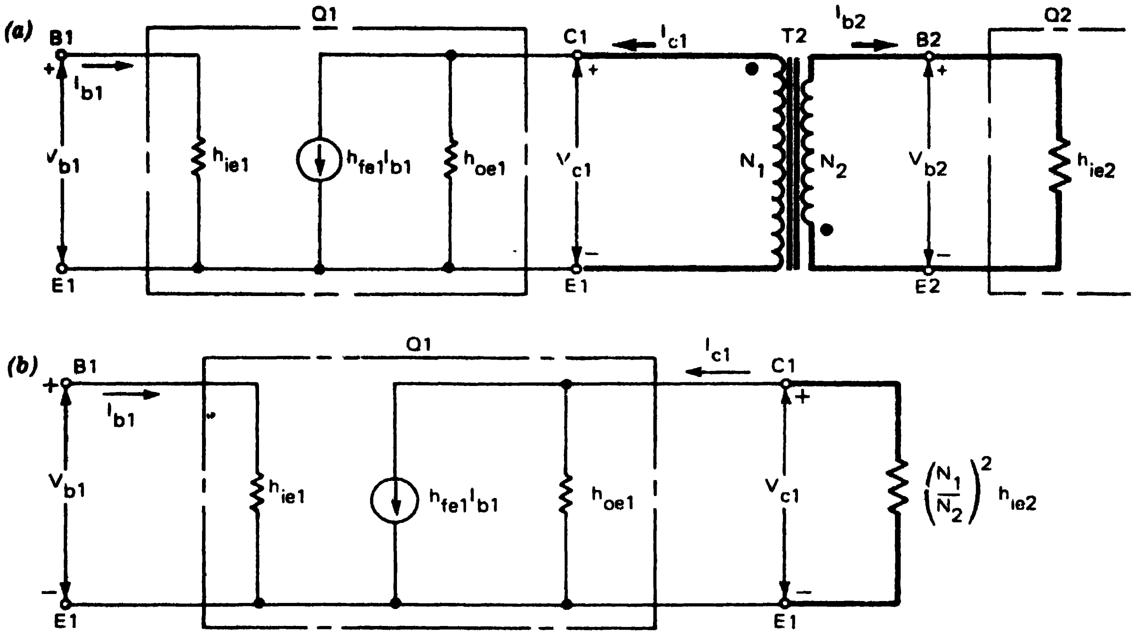

(a)
B1
Q1
C1
I_{c1}
T2
I_{b2}
B2
Q2
I_{b1}
V_{b1}
h_{ie1}
$h_{fe1}I_{b1}$
h_{oe1}
V_{c1}
N_1
N_2
V_{b2}
h_{ie2}
E1
E1
E2
(b)
B1
Q1
I_{c1}
C1
I_{b1}
V_{b1}
h_{ie1}
$h_{fe1}I_{b1}$
h_{oe1}
V_{c1}
$\left(\frac{N_1}{N_2}\right)^2 h_{ie2}$
E1
E1

Fig. 15-7. (a) Approximate equivalent circuit for a transformer-coupled stage. (b) Simplified equivalent circuit showing the reflected impedance $\left(\frac{N_1}{N_2}\right)^2 h_{ie2}$; $h_{ie1} = h_{ie2} = 1$ *K ohm,* $h_{fe1} = 50$, $h_{oe1} = 25 \times 10^{-6}$ *mho.*

impedance is $\frac{1}{h_{oel}}$, we need these to be equal for maximum power transfer. That is,

$$\left(\frac{N_1}{N_2}\right)^2 h_{ie2} = \frac{1}{h_{oel}}. \qquad (15\text{-}15)$$

$$\therefore \frac{N_1}{N_2} = \frac{1}{\sqrt{h_{ie2}\,h_{oel}}} \qquad (15\text{-}16)$$

$$= \frac{1}{\sqrt{1000 \times 25 \times 10^{-6}}} = 6.3.$$

Thus, T2 is a step-down transformer, having a primary-to-secondary turns ratio of 6.3 to 1. Since $\frac{1}{h_{oel}} = 40$ K ohms, the transformer would be specified as matching a 40 K ohms primary to a 1 K ohm secondary. Actually, the output impedance of Q1 is something closer to 50 K ohms if the source resistance, R_s, is taken into account.

To determine the current gain, assume I_{b1} flowing into B1, as shown in Fig. 15-7(a). The transistor current generator is $h_{fel}I_{b1}$, so the collector current, I_{c1}, through the transformer primary will be $\frac{h_{fel}I_{b1}}{2}$ for matched conditions. Since T2 is a voltage step-down transformer, the current on the secondary will be stepped-up by the ratio $\frac{N_1}{N_2}$. Thus, the expression for I_{b2} becomes

$$I_{b2} = \frac{N_1}{N_2} I_{c1} = \frac{N_1}{N_2} \frac{h_{fe1} I_{b1}}{2},$$

and the expression for current gain is

$$A_i = \frac{I_{b2}}{I_{b1}} = \frac{N_1}{N_2} \times \frac{h_{fe1}}{2} \qquad (15\text{-}17)$$

$$= 6.3 \times \frac{50}{2} = 158.$$

This current gain is approximately eight times as large as the stage gain evaluated for the capacitor-coupled amplifier in Section 15-3. This is because there is no current division between load resistance, biasing resistors, and base input. All the collector current magnified by the transformer turns ratio is the base input current to the next stage.

The expression for current gain given by Eq. 15-17 may also be used for the second stage if T3 matches the load resistance R_L to the output resistance of Q2. Similarly, the current gain $\frac{I_{b1}}{I_i}$ will be given by $\frac{N_1}{N_2}$ for T1.

Mid-Frequency Stage Gain, A_v

Let us now obtain expressions for $\frac{V_{c1}}{V_{b1}}$ and the stage gain, $A_v = \frac{V_{b2}}{V_{b1}}$. From Fig. 15-7(b), V_{c1} is given by

$$V_{c1} = -I_{c1} \left(\frac{N_1}{N_2}\right)^2 h_{ie2}, \qquad (15\text{-}18)$$

$$\text{but } I_{c1} = \frac{h_{fe1} I_{b1}}{2}, \text{ and } I_{b1} = \frac{V_{b1}}{h_{ie1}}. \qquad (15\text{-}19)$$

Substituting Eq. 15-19 into 15-18 yields

$$\frac{V_{c1}}{V_{b1}} = \frac{-h_{fe1}}{2} \times \frac{h_{ie2}}{h_{ie1}} \left(\frac{N_1}{N_2}\right)^2, \qquad (15\text{-}20)$$

$$\text{but } \frac{V_{c1}}{V_{b2}} = \frac{-N_1}{N_2}. \qquad (15\text{-}21)$$

which, substituted into Eq. 15-20, gives the expression for the stage voltage gain,

$$A_v = \frac{V_{b2}}{V_{b1}} = \frac{h_{fe1}}{2} \times \frac{h_{ie2}}{h_{ie1}} \times \frac{N_1}{N_2}. \qquad (15\text{-}22)$$

If equal transistors are used, with $h_{ie2} = h_{ie1}$, as in our example, Eq. 15-22 becomes

$$A_v = \frac{N_1}{N_2} \times \frac{h_{fe1}}{2}, \qquad (15\text{-}23)$$

which is the same expression as for the current gain, A_i, = 158.

Note that a large voltage gain results, in spite of the presence of a step-down transformer. This is because the transistor is working into a much higher effective load resistance than in the capacitor-coupled amplifiers in Section 15-3, resulting in a larger voltage gain between base and collector. In the above example, $\frac{V_{c1}}{V_{b1}} = 990$.

For the matched conditions in the above example we would have a stage power gain $A_p = A_v A_i = 158^2 = 25{,}000$ or 44 dB. It should be evident, then, that fewer stages would be necessary to provide a required power gain if transformer coupling were used instead of capacitor coupling. However, this must be balanced against the cost of the transformers (often more expensive than the transistors) and the effect of the transformer on the frequency response.

15-7. LOW-FREQUENCY RESPONSE OF A TRANSFORMER-COUPLED AMPLIFIER

A more accurate equivalent circuit for the transformer at low frequencies includes the primary inductance, L_p, so that the load reflected from the secondary into the primary is shunted by L_p, as shown in Fig. 15-8.

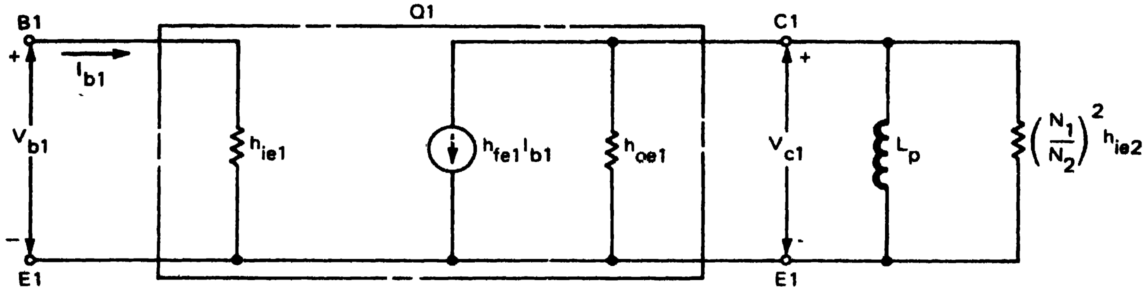


Fig. 15-8. Approximate equivalent circuit for the transformer-coupled stage at low frequencies, showing the primary inductance, L_p.

At lower frequencies, the shunting effect of L_p will reduce the effective load for Q1, with a reduction in V_{c1} and voltage gain. If R represents the parallel combination of $\frac{1}{h_{oe1}}$ and $\left(\frac{N_1}{N_2}\right)^2 h_{ie2}$, the lower 3-dB frequency, f_1, will occur when the reactance of L_p equals R. That is,

$$2\pi f_1 L_p = R, \tag{15-24}$$

or

$$f_1 = \frac{R}{2\pi L_p}, \tag{15-25}$$

where

$$R = \frac{1}{h_{oe1}} \;\Big\|\; \left(\frac{N_1}{N_2}\right)^2 h_{ie2}. \tag{15-26}$$

For the example under consideration, R = 40 K ‖ 40 K = 20 K, so, to obtain a lower 3-dB frequency f_1 = 100 Hz, we would need a transformer whose primary inductance is given by

$$L_p = \frac{R}{2\pi f_1} = \frac{20{,}000}{2\pi 100} = 32 \text{ H}.$$

Values of inductance available in a transformer range from 0.6 mH to 150 H, with the larger inductance obviously involving a larger transformer and more cost. If R is reduced, the necessary value of L_p is also reduced (or f_1 is reduced for a given value of L_p), but this will destroy the maximum power transfer conditions, so A_p will suffer.

15-8. STEP RESPONSE OF A TRANSFORMER-COUPLED AMPLIFIER

Using the expression for sag from Eq. 15-13 and the expression for f_1 from Eq. 15-25, we can find what factors determine the pulse amplifying capability of a transformer-coupled amplifier.

$$\text{Sag} = \frac{V - V_1}{V} = \frac{\pi f_1}{f} = \frac{\pi}{f} \times \frac{R}{2\pi L_p} = \frac{R}{2fL_p}. \tag{15-27}$$

$\frac{1}{2f} = \frac{T}{2}$ = half the period of the pulse and may be designated as t_d, the duration of the pulse. Thus,

$$\text{Sag} = \frac{R}{L_p}\, t_d, \tag{15-28}$$

and the value of L_p necessary to reproduce a square or flat-topped pulse of duration t_d with a certain maximum sag is given by

$$L_p = \frac{Rt_d}{Sag}. \qquad (15\text{-}29)$$

For example, if a transformer-coupled amplifier with R = 20 K ohms is to reproduce a flat-topped pulse of duration 1 msec with a maximum sag of 10%, L_p = 200 H. This is an impractically large value, so that the amplifier would have to be redesigned for a lower value of R, say 1 K ohm, in which case L_p = 10 H. On the other hand, if the pulse duration were only 1 μsec, with R = 1 K ohm, L_p would only have to be 10 mH, so that a much smaller and less expensive transformer could be used.

15-9. SUMMARY

(1) The value of the emitter bypass capacitor in a single-stage CE amplifier must be very large to provide a low value for the lower 3-dB frequency, f_1. Equations for f_1 for different circuit conditions are summarized in Table 15-1.

(2) The electrolyte losses in the emitter bypass capacitor affect the low-frequency response somewhat, but have a significant effect on mid-frequency gain.

(3) Low-frequency response is also affected by the size of the coupling capacitors, as shown by

$$f_1 = \frac{1}{2\pi(R_s + R_i')C_c}$$

where R_i' is the effective input resistance to the amplifier.

(4) The overall mid-frequency voltage gain of an amplifier consisting of cascaded CE stages is the product of the gains of all of the stages, but the overall current gain must take into account the load and biasing resistors, along with the transistor input impedances.

(5) Analysis of cascaded stages generally proceeds from the last stage back towards the input, by determining A_i, R_i and A_v in that order.

(6) The overall low-frequency response, $f_{1(n)}$, for n cascaded stages each having the same value of f_1 is given by

$$f_{1(n)} = \frac{f_1}{\sqrt{2^{\frac{1}{n}} - 1}},$$

which shows that the low-frequency response is poorer than for a single stage.

(7) The ability of an amplifier to handle a square-wave signal is measured by the sag and is related to the amplifier's lower 3-dB frequency, f_1, by

$$Sag = \frac{V - V_1}{V} = \frac{\pi f_1}{f} \times 100\%,$$

where f is the frequency of the square wave. To provide low sag requires a very low value for f_1.

(8) Cascaded transformer-coupled amplifiers designed for maximum power transfer conditions have equal current and voltage gains given by

$$A_i = A_v = \frac{N_1}{N_2} \times \frac{h_{fe}}{2},$$

if equal transistors are used in each stage.

(9) The lower 3-dB frequency, f_1, for a transformer-coupled amplifier is restricted by the transformer's primary inductance, L_p, as given by

$$f_1 = \frac{R}{2\pi L_p},$$

where R is the effective resistive portion of the transistor's load and output resistance in parallel with L_p.

(10) In order for a transformer-coupled amplifier to handle a pulse of duration t_d with less than a given amount of sag, the transformer must have a minimum primary inductance given by

$$L_p = \frac{Rt_d}{Sag}.$$

QUESTIONS

Q15-1. What is the general equation for $\frac{A_{v(LF)}}{A_{v(MF)}}$ as a function of frequency for a CE amplifier with a bypassed resistor in the emitter lead? What does this equation assume regarding C_c?

Q15-2. What is an expression for the lower 3-dB frequency, f_1, for the amplifier in Q15-1? Under what conditions is this expression valid?

Q15-3. When may the amplifier in Q15-1 not have a lower 3-dB frequency due to C_e? Why do you think this happens?

Q15-4. What effect might the emitter bypass capacitor have at mid frequencies? What is this caused by? How does this modify the expression for voltage gain at mid frequencies?

Q15-5. What is the expression for f_1 as a function of C_c if C_e is very large? If the series resistance of C_e is increased, what will be the effect on f_1?

Q15-6. Which do you think usually limits an amplifier's lower 3-dB frequency, C_e or C_c? Why?

Q15-7. When two transistors in the CE connection are cascaded, what does the effective load on the first stage consist of? What effect does this have on the voltage gain of the first stage?

Q15-8. Why is it that the total voltage gain of a cascaded amplifier is the product of the stage gains, but the total current gain is *not* the product of the stage current gains?

Q15-9. The overall voltage gain of a cascaded amplifier is the product of the amplifier's current gain and load resistance divided by the input resistance: true or false? Which equation regarding voltage gain in Chapter 14 is this reminiscent of?

Q15-10. Is the overall value of f_1 for a cascaded amplifier higher or lower than the value of f_1 for any of the individual stages? Is the bandwidth of the overall amplifier wider or narrower than that of one of the stages?

Q15-11. Describe what you understand by the term "sag," drawing a diagram if necessary. How does the lower 3-dB frequency of an amplifier determine the amount of sag in the output when handling a square wave?

Q15-12. What is the principal advantage of using transformers to cascade transistor amplifier stages? How would an interstage transformer differ from an output transformer? How may it differ from a driver transformer?

Q15-13. Why is current gain larger in transformer-coupled stages? Assuming transformers have been chosen for maximum power transfer conditions, what is the expression for a stage's current gain and voltage gain if equal transistors are used in each stage?

Q15-14. What contributes to the drop-off in gain at lower frequencies in a transformer-coupled amplifier? What does this mean if we need a lower value of f_1?

Q15-15. Why is it possible to amplify short-duration pulses in a given transformer-coupled amplifier with less sag than pulses of longer duration? Which equation supports this?

PROBLEMS

P15-1. Consider the CE amplifier in Fig. 15-1(a). $V_{CC} = 12$ V, $R_L = 4.7$ K ohms, $R_e = 1$ K ohm, $R_1 = 15$ K ohms, $R_2 = 2.2$ K ohms, $R_s = 600$ ohms. The transistor has h parameters of: $h_{fe} = 400$, $h_{ie} = 10$ K ohms. Assume $C_e = 50\ \mu F$ with no series resistance. Calculate:

(a) The mid-frequency voltage gain, $\frac{V_o}{V_s}$.

(b) The mid-frequency voltage gain, $\frac{V_o}{V_i}$.

(c) $\frac{(1 + h_{fe})R_e}{R_s + h_{ie}}$ Is this much larger than 1? What does this mean?

(d) f_o and f_p. Is $f_p >> f_o$? What does this mean?

(e) f_1 for $\frac{V_o}{V_s}$.

(f) f_1 for $\frac{V_o}{V_i}$.

P15-2. (a) Calculate the size of bypass capacitor C_e for the amplifier in P15-1 to have a lower 3-dB frequency f_1 of 50 Hz for $\frac{V_o}{V_i}$. (Assume $R_C = 0$.)

(b) What is the mid-frequency voltage gain in (a)?

(c) What is the mid-frequency voltage gain, V_o/V_i, if the capacitor in (a) has a series resistance of 20 ohms?

(d) What will be the value of f_1 with the capacitor in (c)?

P15-3. Consider Eq. 15-2.

(a) If $\frac{(1 + h_{fe})R_e}{R_s + h_{ie}} = 0.4$, express f_p in terms of f_o.

(b) Write Eq. 15-2 as a function of f and f_o

(c) Evaluate $\frac{A_{v(LF)}}{A_{v(MF)}}$ at (1) $f = 1.4f_o$

(2) $f = f_o$

(3) $f = 0$.

Do you see why the expression given in Eq. 15-4 is the condition for no lower 3-dB point due to C_e?

P15-4. Assume the amplifier in P15-1 is modified so that $R_e = 10$ ohms, and that R_1 and R_2 are made large compared with h_{ie}. $C_e = 50\ \mu F$, $h_{ie} = 10$ K ohms, $h_{fe} = 400$, $R_s = 600$ ohms. Calculate:

(a) $\frac{(1 + h_{fe})R_e}{R_s + h_{ie}}$ Is this less than 0.414? What does this mean?

(b) f_o and f_p.

(c) $\frac{A_{v(LF)}}{A_{v(MF)}}$ (Eq. 15-2) at (1) $f = f_p$

(2) $f = f_o$

(3) $f = 0$.

P15-5. Using the values in P15-1, assume C_e is sufficiently large so that f_1 is determined by C_c.

(a) If $C_c = 0.1\ \mu F$, find f_1 where $\frac{V_o}{V_i}$ drops by a factor of 0.707.

(b) Determine the value of C_c to make $f_1 = 50$ Hz if it has zero series resistance.

(c) If C_e alone causes f_1 to be 100 Hz and C_c alone causes f_1 to be 10 Hz, what is the overall lower 3-dB frequency?

P15-6. Two CE stages are cascaded in a similar manner to that in Fig. 15-3 with $R_L = 4.7$ K ohms, $R_e = 1$ K ohm, $C_{e1} = 100\ \mu F$, $C_{e2} = 30\ \mu F$, $C_{c1} = 3\ \mu F$, $C_{c2} = 1\ \mu F$, $R_1 = 15$ K ohms, $R_2 = 2.2$ K ohms, $h_{ie} = 4$ K ohms, $h_{fe} = 100$, and $R_s = 600$ ohms. Calculate:

(a) The overall voltage gain, V_3/V_1.

(b) The overall current gain, I_{c2}/I_1, taking into account all resistances.
(c) The overall power gain in decibels.
(d) The overall value of f_1 for V_3/V_s due to the coupling capacitors alone.
(e) The overall value of f_1 due to the emitter bypass capacitors alone.
(f) The overall value of f_1 due to C_c and C_e.

P15-7. A CE amplifier is providing gain for a square wave.

(a) If the leading edge has an amplitude of 8 V and the falling edge 7 V, what percentage sag does this represent?
(b) If the frequency of the square wave is 50 Hz, what is the lower 3-dB frequency of the amplifier?
(c) If the amplifier's value of f_1 is raised to 100 Hz, what percentage sag would occur in a 350-Hz square wave?
(d) If the leading edge of the square wave in (c) is 8 V, what is the amplitude of the falling edge?
(e) If the amplifier is driven from a source of 600 ohms, the transistor's h_{ie} = 4 K ohms, and h_{fe} = 100, determine the minimum value of C_e to provide the value of f_1 in (c).

P15-8. Consider the amplifier in Fig. 15-6. Q1 and Q2 are identical with h_{ie} = 4 K ohms, h_{fe} = 100, h_{oe} = 50 μmhos. The transformers are to be selected to provide maximum power transfer.

(a) If R_s = 1 K ohm and R_L = 4 ohms, specify the primary and secondary impedances of transformers T1, T2, and T3. Assume the approximate circuit in Fig. 15-7 may be used such that $R_i \approx h_{ie}$. (h_{re} = 0.)
(b) Determine the turns ratio of T1; hence obtain the current gain I_{b1}/I_i.
(c) Obtain the current gain of the first stage, I_{b2}/I_{b1}, using Eq. 15-17.
(d) Obtain the current gain of the second stage, I_L/I_{b2}.
(e) Thus determine the overall current gain, I_L/I_i.
(f) From the turns ratio in (b) obtain the voltage gain V_{b1}/V_{in}.
(g) Obtain the voltage gain V_{b2}/V_{b1}, using Eq. 15-23.
(h) Obtain the voltage gain, $V_L/V_{b2} = N_1/N_2 \times h_{fe2}/2 \times R_L/h_{ie2}$.
(i) Thus, determine the overall voltage gain, V_L/V_{in}.
(j) Determine the overall power gain in decibels.

P15-9. Repeat P15-8(a) using $h_{re} = 7 \times 10^{-4}$ and the dual matching equations (14-15).

P15-10. (a) Determine the minimum primary inductance that T2 must have in P15-8 so that the first stage has an f_1 = 100 Hz for V_{c1}/V_{b1}.
(b) If the primary inductance of T3 is 16 H, determine the lower 3-dB frequency, f_1, for the second stage, V_{c2}/V_{b2}.
(c) What would be the maximum duration of a pulse that the second stage could handle without exceeding a 1% sag in the output?

16

HIGH-FREQUENCY RESPONSE OF THE TRANSISTOR AMPLIFIER

16-0. INTRODUCTION

The high-frequency response of a transistor amplifier, unlike the low-frequency response, depends primarily on the capability of the transistor itself. In this chapter, we examine a hybrid-π model for the transistor, which takes into account the capacitive effects of the PN junctions. This model permits the determination of the CE short-circuit current gain and its dependence upon frequency. As a result, we define the alpha and beta cutoff frequencies and examine what is meant by the gain-bandwidth product, a very important high-frequency characteristic of a transistor. Equations for the CE upper 3-dB frequency, f_2, are developed, which take into account the Miller effect of the capacitance between output and input. And the significance of an amplifier's bandwidth in determining the rise time of a step voltage input is examined in both a capacitor-coupled and a transformer-coupled amplifier.

Finally, because it varies with frequency, the noise figure of a transistor is considered, along with its variation with collector current and source resistance.

16-1. HIGH-FREQUENCY MODEL FOR THE COMMON-EMITTER AMPLIFIER

In order to determine the upper 3-dB frequency, f_2, for the CE amplifier, we must consider the capacitive effects of the transistor junctions. A hybrid-π, or Giacoletto, model which shows the capacitances involved is shown in Fig. 16-1(a).

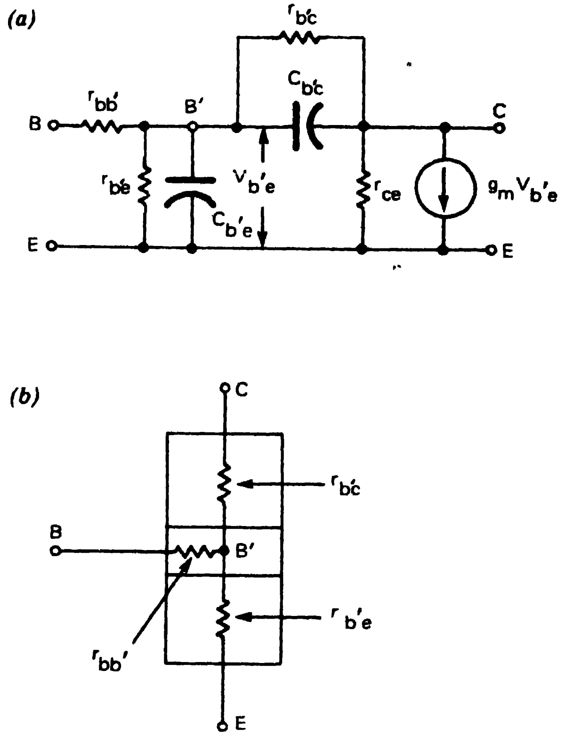


Fig. 16-1. (a) Hybrid-π model for a transistor in the CE connection. (b) Diagram showing the ohmic base-spreading resistance, $r_{bb'}$, and the virtual base, B'.

The Hybrid-π Capacitances

It was shown in Section 5-15 that a forward-biased PN junction exhibited a capacitive effect called the diffusion capacitance. This capacitive effect of the normally forward-biased base-emitter junction of the transistor is shown as $C_{b'e}$ in Fig. 16-1(a). Similarly, it was shown that a reverse-biased PN junction has a transition capacitance, so that the transistor's collector-base junction capacitance may be shown as $C_{b'c}$ in Fig. 16-1(a).

Base-Spreading Resistance

Note that these capacitances have a common termination at B′, not the input base terminal, B. The reason for this may be seen by referring to the diagram in Fig. 16-1(b). Resistance $r_{bb'}$, referred to as the base-spreading resistance, represents the bulk resistance of the base and is the portion of the base through which the base current proper flows. Resistance $r_{b'e}$ is that portion of the base and emitter which may be thought of as being "in series with" the collector junction. This establishes a *virtual* base B′ for the junction capacitances to be connected to instead of B. At low frequencies, these capacitive effects are negligible, so that the input resistance from base-to-emitter with the output shorted is simply $r_{bb'} + r_{b'e}$. But this is the same as h_{ie}, thus:

$$h_{ie} = r_{bb'} + r_{b'e}. \tag{16-1}$$

Resistance $r_{b'c}$ that shunts $C_{b'c}$ is very large, typically 4 M ohms, so that it can normally be omitted. Resistance r_{ce} is the output resistance, and, since $r_{ce} >> R_L$, we can usually omit r_{ce} when we connect a load, R_L, from collector to emitter.

Transistor Transconductance

The mutual conductance or transconductance of a transistor g_m, is defined as follows:

$$g_m = \frac{\Delta I_C}{\Delta V_{B'E}} \text{ at a constant } V_{CE}. \tag{16-2}$$

Obviously $V_{B'E}$ is not a measurable quantity, so that g_m may be approximated by

$$g_m \approx \frac{\Delta I_C}{\Delta V_{BE}},\ V_{CE} \text{ constant}, \tag{16-3}$$

but more accurately,

$$g_m \approx \frac{I_C}{V_T}, \tag{16-4}$$

where, from Section 5-12, $V_T = \frac{T}{11{,}600}$, T in °K.

At room temperature, $V_T = 0.026$ volts, so that g_m at 22°C is given by

$$g_m = \frac{I_C(\text{mA})}{26}. \tag{16-5}$$

Thus, if $I_C = 1$ mA, $g_m \approx 38$ mA/V or 38 millimhos, a value much larger than for a tube.

Note that g_m is required, since, in the high-frequency model of Fig. 16-1(a), the output current generator has a value given by $g_m V_{b'e}$.

16-2. APPROXIMATE CE HIGH-FREQUENCY MODEL WITH A RESISTIVE LOAD

Assuming $r_{b'c}$ is very large and $r_{ce} >> R_L$, we can use the approximate model in Fig. 16-2 to investigate the high-frequency response.

The parameters for the approximate hybrid-π model may be obtained from the following equations. (Numerical values shown are for a 2N2218 silicon transistor with $h_{fe} = 50$, $h_{ie} = 1$ K ohm, $f_T = 300$ MHz, $C_{ob} = 4$ pF, $I_C = 1.5$ mA, T = 22°C.)

$$g_m = \frac{I_C}{V_T} = \frac{I_C(\text{mA})}{26} = 58 \text{ mA/V}.$$

$$r_{b'e} = \frac{h_{fe}}{g_m} = 860 \text{ ohms}.$$

$$r_{bb'} = h_{ie} - r_{b'e} = 140 \text{ ohms}. \tag{16-6}$$

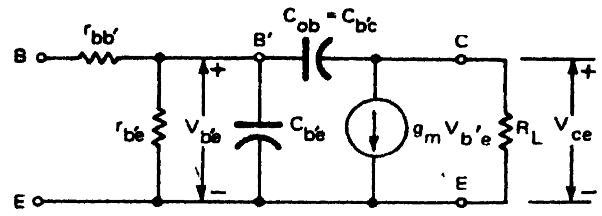


Fig. 16-2. Approximate high-frequency model with resistive load.

$$C_{b'c} = C_{ob} = 4 \text{ pF}.$$

$$C_{b'e} = \frac{g_m}{2\pi f_T} - C_{b'c} = 27 \text{ pF}.$$

$C_{b'c}$ is usually given in the transistor manual as C_{ob}, and is the CB open-circuit output capacitance. $C_{b'e}$ may be given by the manufacturer, or it may be calculated using the equation given. f_T is the frequency at which the CE short-circuit current gain drops to unity, as explained in the next section.

16-3. CE SHORT-CIRCUIT CURRENT GAIN

In order to obtain some idea of a transistor's high-frequency capability and what transistor to choose for a given application, we examine the way in which a transistor's CE short-circuit forward-current transfer ratio or current gain varies with frequency.

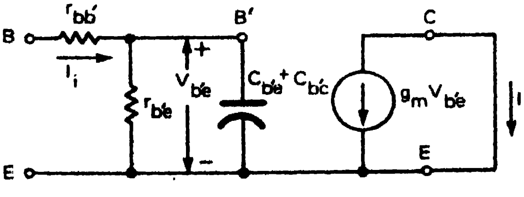


Fig. 16-3. Approximate high-frequency circuit for determination of the short-circuit CE current gain.

When we short circuit R_L in Fig. 16-2, the approximate high-frequency model becomes that shown in Fig. 16-3, where $C_{b'c}$ is now in parallel with $C_{b'e}$.

Using the notation shown in Fig. 16-3, the current gain becomes

$$A_i = \frac{I_L}{I_i} = \frac{-h_{fe}}{1 + j\left(\frac{f}{f_\beta}\right)}, \tag{16-7}$$

where

$$f_\beta = \frac{1}{2\pi r_{b'e}(C_{b'e} + C_{b'c})}, \tag{16-8}$$

and h_{fe} = the short-circuit small-signal forward-current transfer ratio at low frequencies, 1 kHz.

β Cutoff Frequency

The β cutoff frequency f_β, also referred to as f_{hfe}, is the CE short-circuit small-signal forward-current-transfer-ratio cutoff frequency. It is the frequency at which a transistor's CE short-circuit current gain drops 3 dB from its value at low frequencies. Thus, f_β represents the maximum attainable bandwidth for the current-gain of a CE amplifier with a given transistor, although the actual maximum will depend upon circuit conditions, as shown in the next sections. The value of

f_β, for the transistor referred to in Eq. 16-6, is given by

$$f_\beta = \frac{1}{2\pi 860(27 + 4)10^{-12}} = 6 \text{ MHz.}$$

α Cutoff Frequency

A transistor used in the CB connection has a much higher 3-dB frequency. This may be shown by considering the approximate high-frequency circuit of the CB with the output shorted; the expression for the current gain is

$$A_i = \frac{I_L}{I_i} = \frac{-h_{fb}}{1 + j\left(\frac{f}{f_\alpha}\right)},$$

$$\text{where } f_\alpha = \frac{1}{2\pi r_{b'e}(1 + h_{fb})C_{b'e}} \approx \frac{h_{fe}}{2\pi r_{b'e}C_{b'e}}$$

$$= \frac{h_{fe}f_\beta(C_{b'e} + C_{b'c})}{C_{b'e}}. \qquad (16\text{-}9)$$

f_α is the alpha cutoff frequency at which the CB short-circuit small-signal forward-current transfer ratio drops 3 dB from its 1 kHz value. Thus, for the 2N2218, with $h_{fe} = 50$, ($h_{fb} = -0.98$), the value of f_α is

$$f_\alpha = 50 \times 6 \times \frac{(27 + 4)}{27} \text{ MHz} = 344 \text{ MHz.}$$

Thus, a 2N2218 with an $h_{fe} = 50$ at 1 kHz will have its short-circuit current gain drop to 50 × 0.707, or approximately 35, when operated at 6 MHz in the CE connection. However, its value of $h_{fb} = -0.98$ at 1 kHz will not drop to (−0.98) × 0.707, or −0.69, until operated at a frequency of 344 MHz in the CB connection. If the transistor were operated in the CE configuration at 344 MHz, its short-circuit current gain would be only 0.88*.

The frequency capability of a transistor in the CB circuit is evidently superior to the CE, although the latter has more gain to start with. The manner in which the short-circuit current gains for the CE and CB connections vary with frequency is shown in Fig. 16-4.

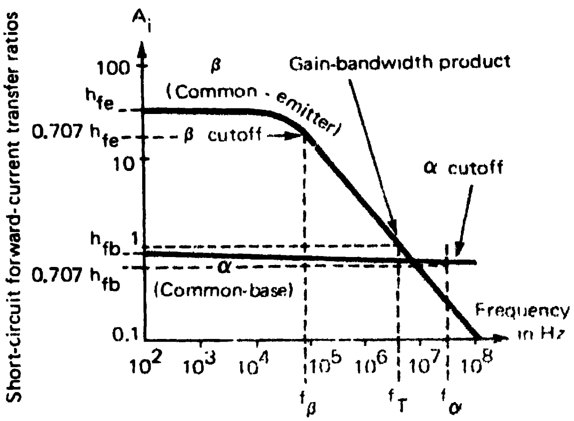


Fig. 16-4. Graph showing how the short-circuit forward-current transfer ratios for the CE and CB vary with frequency. (Typical values are shown for an audio transistor.)

Gain-Bandwidth Product

Although f_α and f_β are useful indications of the high-frequency capability of a transistor, an even more important characteristic is f_T. This is defined as the frequency at which the short-circuit common-emitter current gain has a magnitude of unity. This is shown in Fig. 16-4 and is something less than f_α, but much larger than f_β. For the 2N2218 transistor, $f_T = 300$ MHz, $f_\alpha = 344$ MHz and $f_\beta = 6$ MHz.

---

*This assumes that Eq. 16-7 is valid at this frequency, which is incorrect, since the hybrid-π model is only valid for $f << 3f_T$. However, it does indicate the extreme reduction in CE current gain at very high frequencies compared with the CB.

Accordingly, from Eq. 16-7,

$$|A_i| = \frac{h_{fe}}{\sqrt{1 + \left(\frac{f}{f_\beta}\right)^2}} \tag{16-10}$$

and at $f = f_T$, $|A_i| = 1$, $\left(\frac{f_T}{f_\beta}\right)^2 >> 1$,

we obtain,

$$\frac{f_T}{f_\beta} \approx h_{fe} \text{ or } f_T \approx h_{fe}\, f_\beta. \tag{16-11}$$

That is, f_T is the product of the low-frequency gain, h_{fe}, and the CE bandwidth, f_β. In fact, f_T is also the product of the low-frequency gain, h_{fb}, and the CB bandwidth, f_α. That is, $f_T \approx h_{fb}\, f_\alpha$. (See Problem 16-4.) Values for f_T range from 1 MHz for audio transistors up to 1 GHz for high-frequency transistors. Using the expression for f_β from Eq. 16-8, f_T becomes

$$\begin{aligned} f_T &= h_{fe}\, \frac{1}{2\pi r_{b'e}(C_{b'e} + C_{b'c})} \\ &= \frac{h_{fe}}{r_{b'e}} \times \frac{1}{2\pi(C_{b'e} + C_{b'c})} \\ &= \frac{g_m}{2\pi(C_{b'e} + C_{b'c})}, \end{aligned} \tag{16-12}$$

and the equation for A_i may be rewritten as

$$A_i = \frac{-h_{fe}}{1 + jh_{fe}\left(\frac{f}{f_T}\right)}. \tag{16-13}$$

This shows how a transistor's short-circuit current-gain depends upon the transistor's gain at low frequencies, h_{fe}, and the high-frequency characteristic, f_T.

Since $C_{b'c}$ is usually small compared with $C_{b'e}$, $f_T \approx \frac{g_m}{2\pi C_{b'e}}$ or, $C_{b'e} \approx \frac{g_m}{2\pi f_T}$, as may be seen in Eq. 16-6.

Another meaning may be ascribed to f_T by referring to Eq. 16-10. When $f >> f_\beta$, the denominator is approximately f/f_β, so that $|A_i|\, f \approx f_\beta\, h_{fe} = f_T$. That is, on the straight line portion of the CE current-gain curve in Fig. 16-4, the product of the frequency and the short-circuit current gain is a constant, equal to f_T. Thus, in an amplifier, increased gain may be obtained only at the expense of bandwidth, and vice-versa, for a transistor with a given f_T. This affords a method of obtaining f_T by measurement, since the current gain A_i may be measured at some known frequency f that is lower than f_T but higher than f_β.

Since $f_T \approx h_{fb} f_\alpha$ and $h_{fb} \approx 1$, then $f_T \approx f_\alpha$. In practice, $f_\alpha \approx 1.2 f_T$ for diffusion transistors and $f_\alpha \approx 2 f_T$ in drift transistors.

16-4. HIGH-FREQUENCY CURRENT GAIN WITH A RESISTIVE LOAD

It should not be inferred from the above that the upper 3-dB frequency, f_2, of a CE amplifier is simply

$f_\beta = \frac{f_T}{h_{fe}}$, since this is the short-circuit-current-gain bandwidth. When the transistor works into a resistive load, we must return to the approximate equivalent circuit of Fig. 16-2, shown for reference in Fig. 16-5(a). Fig. 16-5(b) shows a further simplification of the high-frequency model, where the capacitance from B′ to C has been replaced by a capacitor at the input having a value of $C_{b'c}(1 + g_m R_L)$.

This is a result of Miller's theorem, which states that if an impedance, Z, is connected between the input and output terminals of a

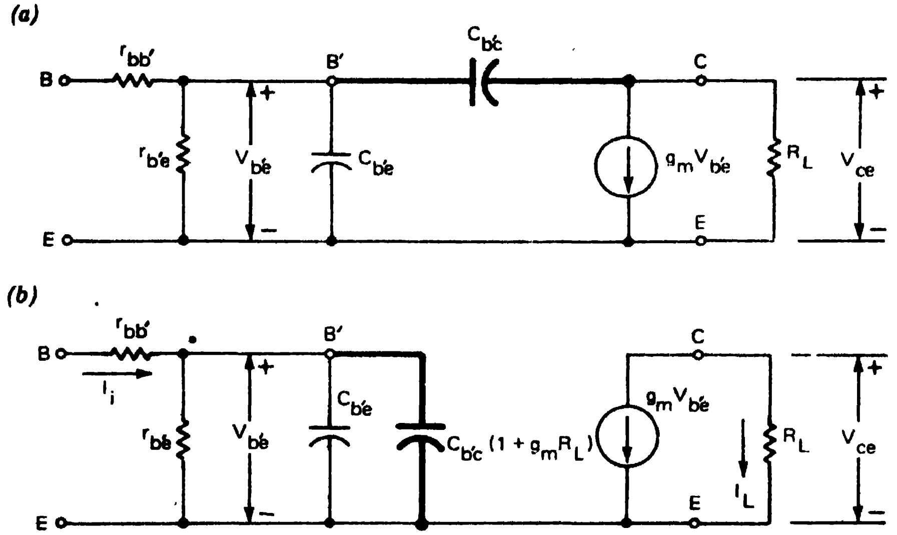


Fig. 16-5. (a) Approximate high-frequency model with resistive load. (b) Approximate model showing an additional capacitance at the input due to the Miller effect.

network between which there is a voltage gain, K, the same effect can be had by removing Z and connecting an impedance, Z_i, at the input equal to $\frac{Z}{1-K}$ and an impedance, Z_o, across the output equal to $\frac{ZK}{K-1}$. See Fig. 16-6.

With reference to Fig. 16-5(b), the voltage gain, $K = \frac{V_{ce}}{V_{b'e}} = -g_m R_L$. Thus, $1 - K = 1 + g_m R_L$, and since Z_i in Fig. 16-6(b) is *smaller* by a factor of $(1 - K)$, the input capacitance in Fig. 16-5(b) is *larger* by an amount of $1 + g_m R_L$. Thus the total input capacitance between B' and E is given by

$$C = C_{b'e} + (1 + g_m R_L)C_{b'c}. \qquad (16\text{-}14)$$

This represents a considerable increase in capacitance such that, if $R_L = 2$ K ohms for

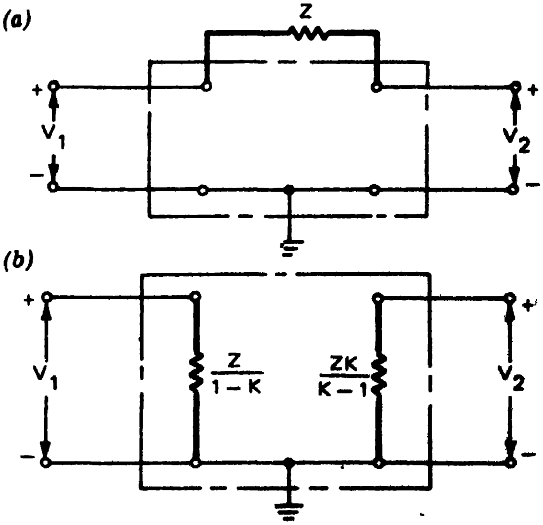


Fig. 16-6. (a) Circuit with impedance between input and output. $K = V_2/V_1$. (b) Showing the same effective circuit due to Miller's theorem.

example, with the parameters used for the 2N2218, C becomes

$$C = 27 + (1 + 58 \times 10^{-3} \times 2 \times 10^{3})4 \approx 500\,\text{pF}.$$

Since the circuit of Fig. 16-5(b) is essentially the same as that of Fig. 16-3 with an additional capacitance (and a load resistance), we may use Eq. 16-8 with the capacitance replaced by C to obtain the upper 3-dB frequency, f_2, for $\frac{I_L}{I_i}$:

$$f_2 = \frac{1}{2\pi r_{b'e} C}. \qquad (16\text{-}15)$$

For the example under consideration, with $r_{b'e} = 860$ ohms, f_2 is given by

$$f_2 = \frac{1}{2\pi 860 \times 500 \times 10^{-12}}\,\text{Hz} = 370\ \text{kHz}.$$

This is appreciably lower than the 6 MHz calculated for f_β, although it may be increased by a reduction in R_L, which would reduce C.

It should be noted that the Miller output capacitance, having an impedance $= \frac{ZK}{K-1}$ in Fig. 16-6(b), was not shown in the output of Fig. 16-5(b). This is because, with K >> 1, the value of output capacitance is approximately equal to $C_{b'c}$. This small value of capacitance (4 pF) in parallel with a low resistance (R_L = 2 K ohms typically) has a negligible time constant compared with the input time constant and can be ignored unless working into a capacitive load.

Effect of Source Resistance on Frequency Response

The mid-frequency voltage gain of a single-stage CE amplifier driven by a voltage source in series with a resistance R_s as shown in Fig. 16-7, is given by the equations in Table 15-1. However, the 3-dB frequency f_2, for current gain, $\frac{I_L}{I}$*, and voltage gain, $\frac{V_o}{V_s}$, is now given by

$$f_2 = \frac{1}{2\pi R'C}, \qquad (16\text{-}16)$$

where R′ is the parallel combination of (R_s + $r_{bb'}$) and $r_{b'e}$.

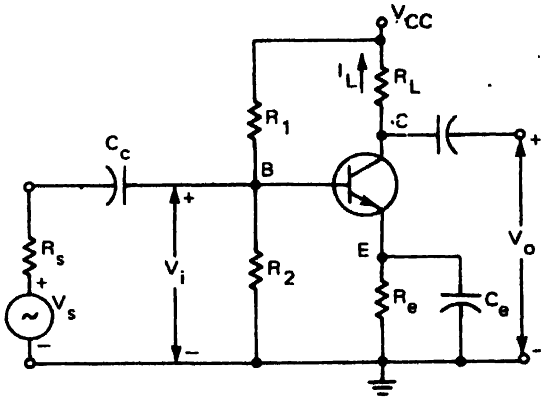


Fig. 16-7. Single-stage CE amplifier to be analyzed for high-frequency response.

That is,

$$R' = \frac{(R_s + r_{bb'})r_{b'e}}{R_s + h_{ie}}, \qquad (16\text{-}17)$$

and C is given by Eq. 16-14. The value for R′ assumes $R = R_1 \| R_2 >> R_s$. If this is not the case, then

$$R' = \frac{(R_s' + r_{bb'})r_{b'e}}{R_s' + h_{ie}}, \text{ where } R_s' = R_s \| R. \qquad (16\text{-}18)$$

---

*Note that I is the magnitude of the current generator when the source is represented by its Norton equivalent.

Using R_s = 600 ohms and the values in Eq. 16-6, we obtain R' = 400 ohms (from 16-17) and f_2 = 800 kHz. If, in addition, $R = R_1 \| R_2$ = 1 K ohm, then R' = 320 ohms (from Eq. 16-18) and $f_2 \approx$ 1 MHz. Thus, the influence of the source and biasing resistors is very strong in determining f_2, with a reduction in the effective source resistance resulting in an increase in f_2.

16-5. HIGH-FREQUENCY RESPONSE OF CASCADED CE STAGES

We can now return to the determination of the overall high-frequency response of the cascaded CE stages shown in Fig. 16-8. This is done by first obtaining f_2 for each stage, but the values to use for R_L and R' in Eq. 16-14 and 16-16 depend upon the stage under consideration, as shown in Example 16-1.

Example 16-1. Using the values shown in Fig. 16-8, determine the value of f_2 for each stage and estimate the overall bandwidth of the amplifier in Fig. 16-8. That is, find f_2 for V_o/V_s.

Solution

Third Stage

R_{L3} = 2 K ohms, C = 500 pF (from Eq. 16-14) and R' in Eq. 16-16 is given by Eq. 16-18, where $R_s = R_{L2}$, the effective output resistance of the second stage.

$$R_s{}' = R_s \| R = R_{L2} \| R = 2\ \text{K} \| 1\ \text{K} = 667\ \text{ohms.}$$

$$R' = \frac{(R_s{}' + r_{bb'})}{R_s{}' + h_{ie}}\, r_{b'e} = \frac{(667 + 140)}{667 + 1000}\, 860$$

$$= 416\ \text{ohms.}$$

$$f_2 = \frac{1}{2\pi R'C} = \frac{1}{2\pi 416 \times 500 \times 10^{-12}}$$

$$= 765\ \text{kHz.}$$

Second Stage

The effective value of load resistance for the second stage, $R_{L2}{}'$, is the parallel combination of R_{L2}, R and h_{ie}.

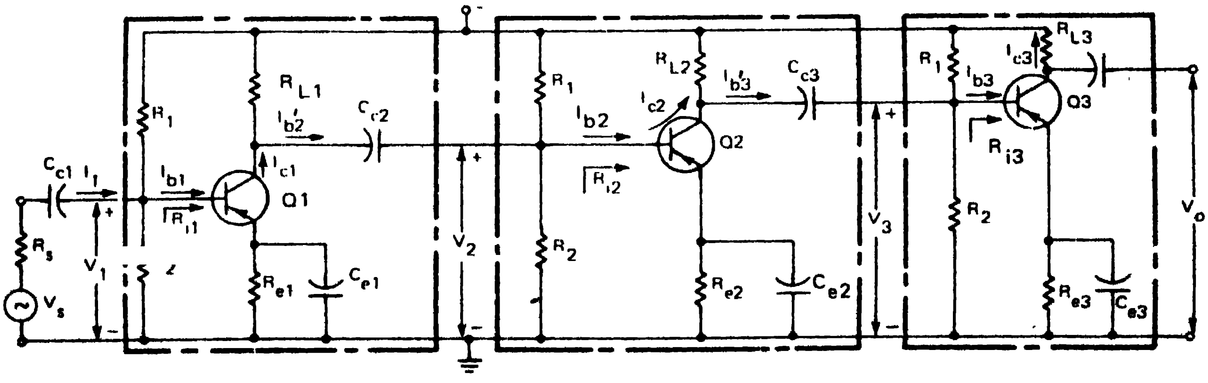


Fig. 16-8. Cascaded CE stages to be examined for high-frequency response. $R_{L1} = R_{L2} = R_{L3}$ *= 2 K ohms,* h_{fe} *= 50,* h_{ie} *= 1 K ohm,* $r_{b'e}$ *= 860 ohms,* $r_{bb'}$ *= 140 ohms,* $C_{b'c}$ *= 4 pF,* $C_{b'e}$ *= 31 pF,* $R = R_1 \| R_2$ *= 1 K ohm,* R_s *= 600 ohms,* g_m *= 58 mmhos.*

$$R_{L2}' = R_{L2} \| R \| h_{ie} = 2\text{ K} \| 1\text{ K} \| 1\text{ K}$$

$$= 400 \text{ ohms.}$$

$$\therefore C = C_{b'e} + (1 + g_m R_{L2}')C_{b'c}$$

$$= 31 + (1 + 58 \times 10^{-3} \times 400)4 = 128 \text{ pF.}$$

R′ for the second stage is the same as for the third stage, 416 ohms.

$$\therefore f_2 = \frac{1}{2\pi R'C} = \frac{1}{2\pi 416 \times 128 \times 10^{-12}}$$

$$= 3 \text{ MHz.}$$

First Stage

The effective load for the first stage, R_{L1}', is the same as $R_{L2}' = 400$ ohms, so that $C = 128$ pF. R′ for the first stage is given by Eq. 16-18 with $R_s = 600$ ohms, $R = 1$ K ohm, so that $R' = 320$ ohms.

$$f_2 = \frac{1}{2\pi R'C} = \frac{1}{2\pi 320 \times 128 \times 10^{-12}}$$

$$= 3.9 \text{ MHz.}$$

Overall Amplifier

The overall value of the bandwidth, f_2, for the cascaded stages, V_o/V_s, will evidently be limited by the value of $f_2 = 765$ kHz for the third stage, since it is approximately one-fourth to one-fifth the value of the other stages. Had the value of f_2 for each stage been the same, the overall value for n cascaded stages would be given by

$$f_{2(n)} = f_2\sqrt{2^{\frac{1}{n}} - 1} \qquad (16\text{-}19)$$

The bandwidth, f_2, of the above amplifier could have been increased by using a smaller value of R_L, especially for the last stage, but the voltage gain would have suffered as a consequence.

16-6. AMPLIFIER HIGH-FREQUENCY RESPONSE TO A SQUARE WAVE

The significance of the bandwidth of an amplifier may be appreciated by considering the response of an amplifier when subjected to a step voltage input as shown in Fig. 16-9(b). The amplifier may be represented by the simple low-pass RC circuit in Fig. 16-9(a).

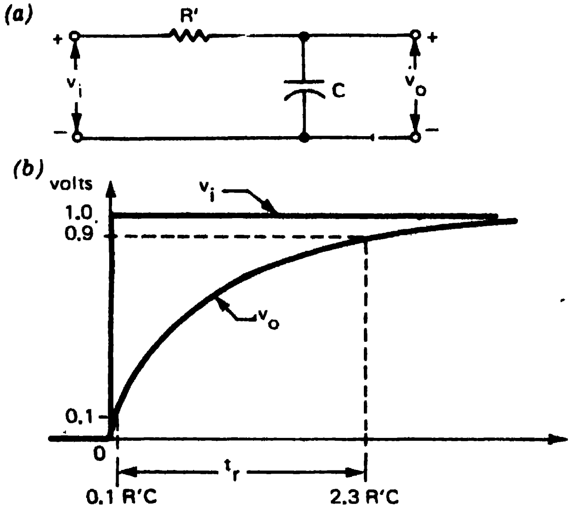


Fig. 16-9. (a) Simplified high-frequency model of an amplifier for rise time considerations.
(b) Response of the RC circuit to a step voltage input.

Assume a 1-volt step input is applied for v_i. The charging of C will result in an exponentially rising voltage, v_o, given by

$$v_o = 1\left(1 - e^{\frac{-t}{R'C}}\right) \qquad (16\text{-}20)$$

The rise time, t_r, is defined as the time for the output to change from 0.1 of its final value to 0.9 of its final value. These times may be obtained from Eq. 16-20 as 0.1R′C and 2.3R′C, respectively, so that t_r is the difference equal to 2.2R′C. That is,

$$t_r = 2.2R'C = \frac{2.2}{2\pi f_2} = \frac{0.35}{f_2}, \qquad (16\text{-}21)$$

where use was made of Eq. 16-16, $f_2 = \frac{1}{2\pi R'C}$.

This indicates that the product of the rise time and the amplifier's bandwidth is a constant, equal to 0.35. Thus, if an oscilloscope is to reproduce a rising edge with a rise time of only 0.35 μsec, a vertical amplifier having a 1 MHz bandwidth is required. For a smaller rise time, such as 350 psec (350×10^{-12} sec) the necessary bandwidth is 1 GHz (1×10^9 Hz), which requires transistors with very low values of junction capacitance. These are not unusual values in high quality modern oscilloscopes, although the increase in bandwidth is usually accomplished by negative feedback methods, a topic covered in Chapter 17.

16-7. HIGH-FREQUENCY RESPONSE OF A TRANSFORMER-COUPLED AMPLIFIER

In Section 15-7, we found an expression for the lower 3-dB frequency, f_1, for a CE stage of a transformer-coupled amplifier of the type shown in Fig. 16-10. This involved the consideration of the transformer's primary inductance, L_p. The shunting effect of this at mid-to-high-frequencies is negligible, but the *series* leakage inductances and distributed capacitances in both primary and secondary now become significant. These are shown in the high-frequency equivalent circuit for a transformer-coupled amplifier in Fig. 16-11(a).

Since the high-frequency capability of the amplifier will normally be limited by the transformer's characteristics and not the transistor's, we can represent the transistor by its approximate low-frequency model. Thus, V_g and R_g in Fig. 16-11 represent either the input source or the equivalent output from the previous transistor. R_L may be the output load or the input to the next stage.

When the resistance, inductance, and capacitance on the secondary side are reflected into the primary, with

$$L_1 = L_{1p} + \left(\frac{N_1}{N_2}\right)^2 L_{1s} \text{ and } C = C_p + \left(\frac{N_2}{N_1}\right)^2 C_s, \qquad (16\text{-}22)$$

we obtain the simplified circuit shown in Fig. 16-11(b). We can see that the circuit will have a series-resonant effect at high frequencies, but because of the usually large value of R_g, the overall Q of the circuit will be low — in the region of 0.5 to 2. Thus the resonant effect is not pronounced, but the resonant frequency, f_o, where $X_C = X_{L1}$, will give an indication of the upper 3-dB frequency, f_2. That is, f_2 is usually somewhat higher than f_o, which is given by

$$f_o = \frac{1}{2\pi\sqrt{L_1 C}}. \qquad (16\text{-}23)$$

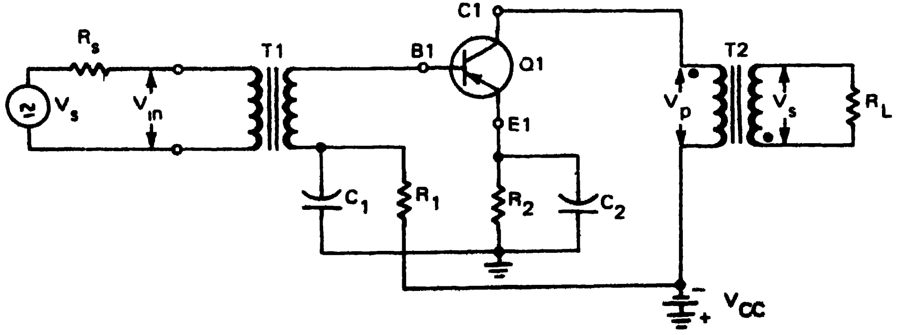


Fig. 16-10. Transformer-coupled amplifier to be analyzed for high-frequency response.

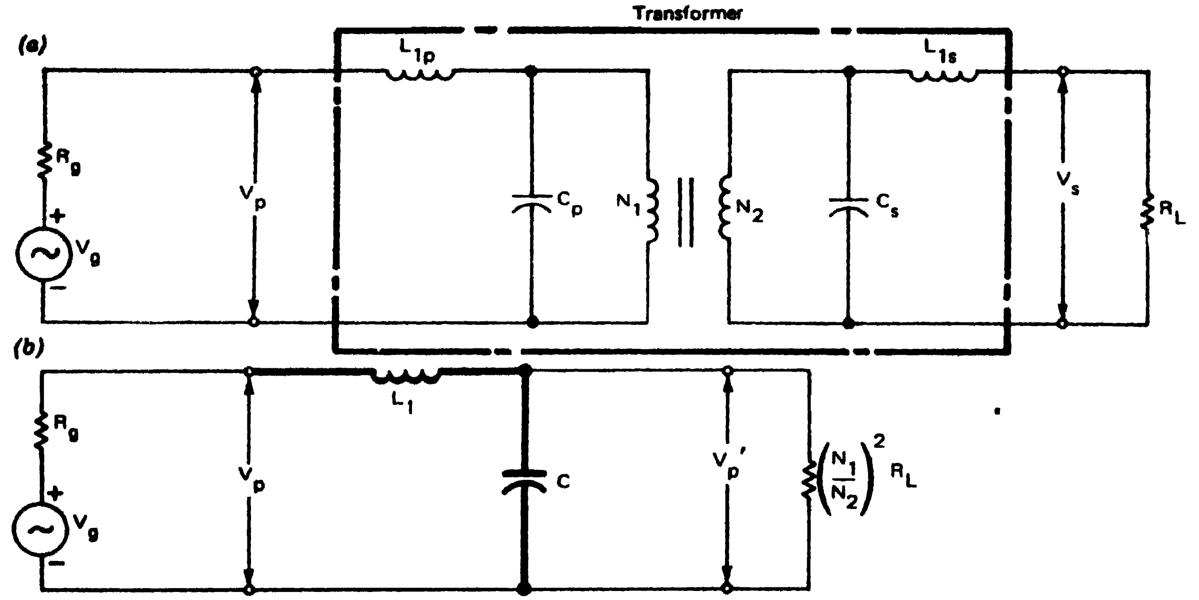


Fig. 16-11. (a) High-frequency equivalent circuit for a transformer-coupled amplifier. (b) Simplified equivalent circuit with all components referred to the primary side of the transformer.

The variation of the amplifier's voltage gain with frequency is shown in Fig. 16-12 for different values of Q.

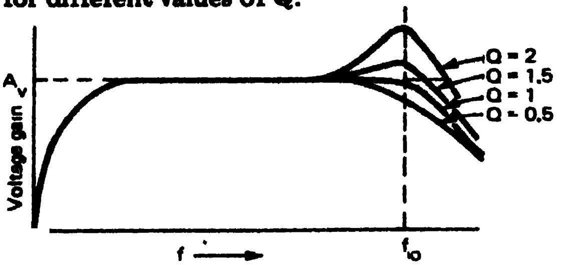


Fig. 16-12. Graph showing how the voltage gain varies with the effective Q of the resonant circuit. (From Semiconductor Devices and Circuits, *p. 204, by Alley and Atwood, Copyright © 1971 John Wiley & Sons, Inc. Used by permission of John Wiley & Sons, Inc.).*

For example, if L_1 = 0.1 H and C = 400 pF, f_o = 25 kHz, and f_2 will be in this neighborhood, with its precise value depending upon the Q of the circuit. However, the value of f_o may be used to predict the effect of the transformer's characteristics on the rise-time capability when handling fast rising pulses, as in switching applications.

This is done by referring to the curves in Fig. 16-13, where the response to a step input of voltage is shown as a function of f_o for a fixed value of Q = 1. The higher the value of f_o (bandwidth), the smaller the rise time. Thus, a transformer with very low values of leakage inductance and distributed capacitance is required for low-rise-time applications.

Example 16-2. A pulse transformer designed to couple fast-rising pulses to the gate of a silicon controlled rectifier (SCR) is advertised as having a maximum series inductance L_1 = 1500 μH and a distributed capacitance C = 400 pF. Assuming the load conditions are

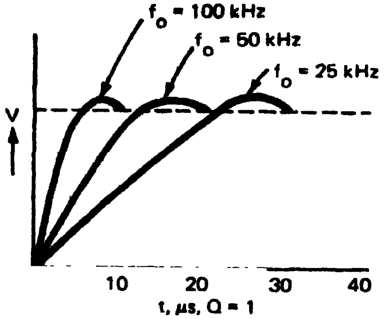


Fig. 16-13. Curves showing how the step response varies with f_o; Q = 1. (From Semiconductor Devices and Circuits, *p. 205, by Alley and Atwood, Copyright © 1971 John Wiley & Sons, Inc. Used by permission of John Wiley & Sons, Inc.).*

such as to provide an effective Q of 1, estimate the rise-time capability of this transformer.

Solution

$$f_o = \frac{1}{2\pi\sqrt{L_1 C}}$$

$$= \frac{1}{2\pi\sqrt{1500 \times 10^{-6} \times 400 \times 10^{-12}}}$$

$\approx$ 200 kHz.

From Fig. 16-13, when f_o = 100 kHz the shortest time to reach the maximum value is approximately 5 μsec, so that for f_o = 200 kHz the time would be in the region of 2.5 μsec. But the rise time is the time to go from 10% to 90% of the final value, so an estimate of the rise time may be 2 μsec.

An alternative approximate solution would be to assume $f_2 \approx f_o$ and use Eq. 16-21.

$$t_r = \frac{0.35}{f_2} = \frac{0.35}{200 \times 10^3} = 1.75\ \mu\text{sec}.$$

16-8. TRANSISTOR NOISE

In Section 8-16, we defined the noise figure, NF, of an amplifier or active device as follows:

$$\text{NF} = 10 \log \frac{\frac{S_i}{N_i}}{\frac{S_o}{N_o}} \qquad (16\text{-}24)$$

where S and N are the signal and noise powers at the input and output. A perfectly quiet device would have a 0 dB noise figure.

The noise contributed by an amplifier or active device comes from a number of sources, including thermal noise, shot noise, and partition noise. See Appendix D, Section 10. All of these are present in transistors, but in specifying transistor noise the frequency, operating point, and source resistance must be given.

If the noise from a device is measured after it passes through a filter which has a very narrow bandpass (a few Hertz), the noise figure is referred to as *spot noise* NF and is the more usual figure specified for a transistor. If the measurement is made with a filter that has a wide bandwidth (usually 15.7 kHz with a lower 3-dB frequency f_1 = 10 Hz), the noise figure is referred to as *broadband noise* or the *average noise* figure, $\overline{\text{NF}}$. This is the figure quoted for transistors that are especially designed for low-level, low-noise applications.

The manner in which the transistor spot-noise figure varies with the source resistance, R_G (R_s), at the input to the transistor, is shown in Fig. 16-14(a) for a TIS37 epitaxial planar silicon transistor.

Evidently the optimum value of R_G is 600 ohms, a value that we have used consistently in the last few chapters. However, the minimum noise point is not always so well defined.

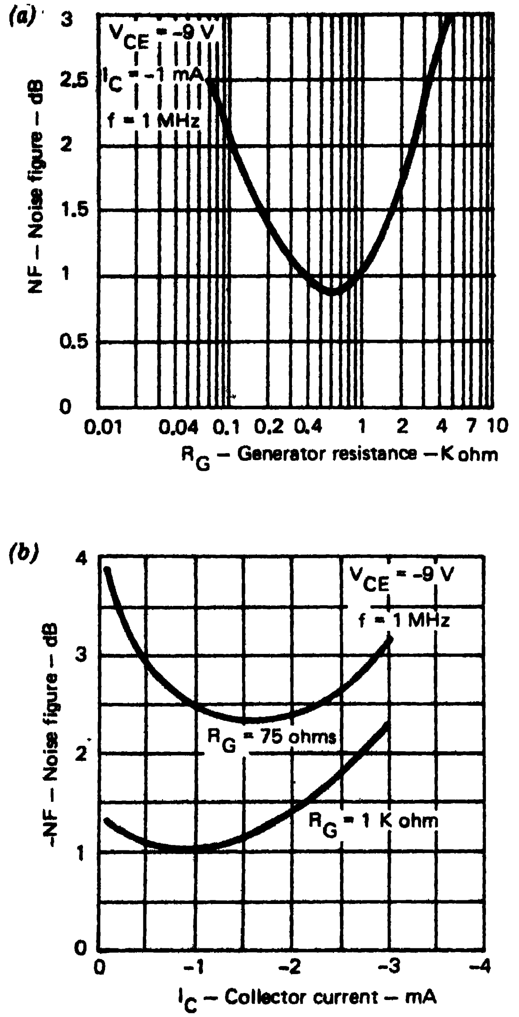


Fig. 16-14. (a) Curve showing the variation of spot noise figure with source or generator resistance for a TIS37 silicon transistor. (b) Curves showing how the noise figure varies with collector current and source resistance. (Courtesy of Texas Instruments, Incorporated.)

The variation of NF with collector current and source resistance is indicated in Fig. 16-14(b).

All of the information in Fig. 16-14 was determined at a frequency of 1 MHz, since NF is also dependent upon frequency, as can be seen by referring to Fig. 16-15.

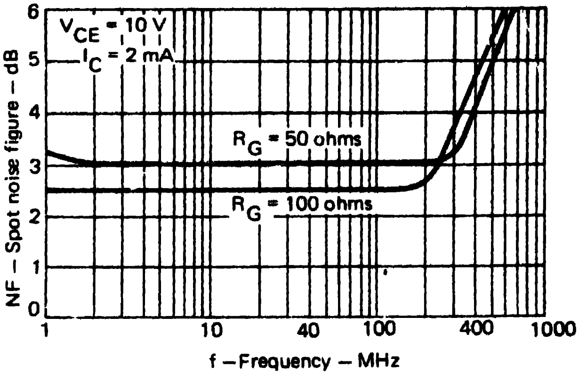


Fig. 16-15. Curve showing how the spot noise figure varies with frequency for a 2N4252 silicon transistor. (Courtesy of Texas Instruments, Incorporated.)

It should be noted that this curve is for a different transistor from that used in Fig. 16-14, but it is typical of the shape; only the frequencies would have to be changed. For frequencies below 1 MHz in Fig. 16-15, the curves increase quite sharply, as they do at high frequencies. Over a large range of midband frequencies, the noise is constant and is referred to as white noise. It is due primarily to thermal and shot noise.

When an amplifier consists of cascaded stages, the noise of the first stage is amplified by the other stages, so it is especially important to use low-noise devices in the earlier stages. The overall effect of noise in an amplifier is heard as a hiss or crackle when the volume of an audio amplifier is turned to maximum with no signal applied. Or, it appears as snow on a TV screen when a weak station is being received. Proper design of the amplifier, which includes judicious selection of active components, is necessary to provide high gain for weak signals from radio, television, and radar, without masking out the signal itself due to noise.

16-9. SUMMARY

(1) The hybrid-π model of a transistor at high frequencies includes the capacitive effects of the PN junctions and involves a base-spreading resistance that creates a virtual base.

(2) The transistor's high-frequency parameters are given in Eq. 16-6 and include the transistor's transconductance.

(3) The transistor's short-circuit current gain varies with frequency and is characterized by the alpha-cutoff and beta-cutoff frequencies, where the current gain drops by 3 dB from the value at mid frequencies in a CB and CE connection, respectively.

(4) An important high-frequency characteristic of a transistor is f_T, the gain-bandwidth product, and is defined to be the frequency at which a common emitter's short-circuit current gain drops to unity.

(5) When a transistor is operated at high frequencies with a resistive load, the collector-to-base junction capacitance appears as an enlarged value of capacitance at the input, thus reducing voltage and current gain below the values expected at the alpha and beta cut-

(6) Cascaded CE stages operating at high frequency have a reduced overall bandwidth given by

$$f_{2(n)} = f_2\sqrt{2^{\frac{1}{n}} - 1}\,,$$

where n is the number of stages, and f_2 is the bandwidth of each stage.

(7) The ability of an amplifier to reproduce a square wave input with a low rise time, t_r, is directly proportional to the bandwidth, f_2. That is, $t_r = 0.35/f_2$.

(8) The high-frequency response of a transformer-coupled amplifier is usually limited entirely by the transformer's leakage inductance and distributed capacitance, which creates a series resonant effect.

(9) Transistor noise is usually measured as a *spot noise* figure, NF, or as the *average noise* figure, $\overline{\text{NF}}$, with typical values between 0.5 and 5 dB depending upon the source resistance, collector current, and frequency.

QUESTIONS

Q16-1. Describe what you understand by the term "virtual base" as used in Fig. 16-1.

Q16-2. Describe how you would determine experimentally an approximate value for a transistor's transconductance, g_m.

Q16-3. How would you expect g_m to vary with temperature? Why?

Q16-4. Distinguish between: f_α and f_β; f_{hfb} and f_{hfe}; f_T and f_α.

Q16-5. Explain how f_T supports the fact that a given transistor's bandwidth may be increased only at the expense of gain in either a CE or CB. How would you determine f_T experimentally?

Q16-6. What is the Miller effect? When is it a factor to take into consideration? What effect does it have on high frequency gain? Why?

Q16-7. Given equal values for R_L in Fig. 16-8, why does the last stage have the lowest value of f_2 of all three stages?

Q16-8. Why can f_2 normally be considered as the bandwidth of an amplifier? When can it not? If three stages each have the same value of f_2, is the overall value of f_2 higher or lower than for one stage?

Q16-9. What is meant by the term "rise time"? How would you measure an amplifier's rise time? What significance does the number 0.35 have? If we want an amplifier with a very low rise time, what must be true about the amplifier's bandwidth?

Q16-10. What is the high-frequency response of a transformer-coupled CE amplifier usually determined by? What causes the possible rise in a transformer-coupled amplifier's frequency response curve at high frequencies? Under what conditions will this not happen?

Q16-11. How does the rise time capability of a transformer-coupled amplifier vary with the resonant frequency, f_o? If Q is increased, what do you think will happen to the rise time? Are there any drawbacks to this?

Q16-12. What would be the noise figure of a perfectly "quiet" active device?

Q16-13. Differentiate between spot noise and broadband noise.

Q16-14. What three factors does spot noise depend upon in an amplifier?

Q16-15. What effect does noise have in an amplifier used in audio or video applications?

PROBLEMS

P16-1. A 40397 silicon transistor used in video applications is connected in the CE configuration with its collector-to-emitter voltage held constant at 1 V. A 70 mV change in the base-emitter voltage causes a change in collector current from 1 mA to 10 mA.

(a) What is the approximate value of the transistor's g_m?
(b) Evaluate g_m at I_C = 1 mA, 4 mA and 6 mA, assuming room temperature and using Eq. 16-5.
(c) If the temperature increased, what factors would have to be taken into consideration to determine the new values of g_m?

P16-2. Given that the 40397 transistor has h_{ie} = 6 K ohms and h_{fe} = 224 at I_C = 1 mA, with f_T = 80 MHz and $C_{b'c}$ = 12 pF. Determine: (a) g_m (b) $r_{b'e}$ (c) $r_{bb'}$ (d) $C_{b'e}$ at room temperature and a collector current of 1 mA. Use Eq. 16-6. Draw the approximate small-signal high-frequency model of the transistor showing numerical values.

Note: In the following problems for the 40397 transistor, use g_m = 38 mmhos, $r_{b'e}$ = 5.9 K ohms, h_{ie} = 6 K ohms, $r_{bb'}$ = 100 ohms, $C_{b'c}$ = 12 pF, $C_{b'e}$ = 63 pF, f_T = 80 MHz and h_{fe} = 224 at 1 kHz.

P16-3. (a) Calculate the α and β cutoff frequencies for the 40397.

(b) What would be the 40397's CE short-circuit current gain at frequencies of:

(1) f_β, (2) f_T, (3) f_α?

(c) Sketch a graph showing how the 40397's CE short-circuit forward-current transfer ratio varies with frequency from 1 kHz to 100 MHz.

P16-4. Use the first expression for f_α in Eq. 16-9 and the expression for f_β in Eq. 16-8 with the assumption that $C_{b'c} << C_{b'e}$ to show that $f_T \approx h_{fb}\, f_\alpha$ given that $f_T \approx h_{fe}\, f_\beta$.

P16-5. A transistor's short-circuit current gain is measured to be 25 at a frequency of 2 MHz. If the transistor's f_β = 200 kHz, determine:

(a) The current-gain-bandwidth product, f_T.
(b) The transistor's h_{fe} at low frequencies.
(c) The short-circuit current gain at 10 MHz and 100 MHz.

P16-6. A 40397 transistor is used in the circuit of Fig. 16-7, with R_L = 1 K ohm and R_s = 600 ohms. Assuming $R_1 || R_2 >> R_s$ and

$R_1 \| R_2 >> R_i$ calculate:

(a) The mid-frequency voltage gain, V_o/V_s.

(b) The upper 3-dB frequency, f_2, for V_o/V_s, taking into account the Miller effect.

P16-7. (a) Repeat P16-6 using $R_1 \| R_2$ = 1 K ohm.

(b) Repeat P16-6 using $R_1 \| R_2$ = 1 K ohm and R_L = 2 K ohms.

P16-8. Three 40397 transistors are used in a cascaded circuit as in Fig. 16-8. Determine the upper 3-dB frequency for each stage and estimate the overall 3-dB frequency, f_2, for V_o/V_s if: R_{L3} = 1 K ohm, R_{L2} = 2.2 K ohms and R_{L1} = 3.3 K ohms. Use $R_1 \| R_2$ = 1 K ohm and R_s = 600 ohms.

P16-9. (a) Repeat P16-8 with $R_{L1} = R_{L2} = R_{L3}$ = 1 K ohm.

(b) Repeat P16-8 with $R_{L1} = R_{L2} = R_{L3}$ = 3.3 K ohms.

(c) In Problem 16-9(a), what should be the value of R_{L3} to make f_2 for the third stage the same as f_2 for the first stage, if $R_{L1} = R_{L2}$ = 1 K ohm?

P16-10. If four cascaded stages must provide an overall bandwidth, f_2, of 1 MHz, what must be the minimum bandwidth of each, assuming all stages have the same f_2?

P16-11. An amplifier must reproduce a square wave with a rise time not to exceed 0.2 μsec.

(a) What must be the minimum upper 3-dB frequency to provide this response?

(b) If the amplifier in Fig. 16-7 is to be used with a 40397 transistor, determine the maximum value of R_L that can be used if R_s = 600 ohms, $R_1 \| R_2$ = 1 K ohm.

(c) What is the gain provided in (b)?

(d) If R_L were double that calculated in (b), what would be the rise time for a step voltage?

P16-12. A transformer-coupled amplifier circuit like that in Fig. 16-10 has an output transformer T2 with the following characteristics: $N_1 : N_2 = 10:1$; L_{1p} = 40 mH, L_{1s} = 200 μH, C_p = 150 pF, C_s = 0.005 μF.

(a) What will the upper 3-dB frequency of the amplifier be limited to (approximately) due to T2?

(b) If the effective output resistance of the transistor is 20 K ohms, what will be the Q of the output circuit at f_o? Will there be a resonant rise in the frequency response around f_o?

(c) What will be the approximate rise time for a square wave signal applied to the above amplifier, based solely on the transistor and T2?

17

NEGATIVE FEEDBACK IN TRANSISTOR AMPLIFIERS

17-0. INTRODUCTION

Methods to stabilize the dc operating point of a transistor amplifier have been investigated. These involved the use of negative feedback as far as dc was concerned to keep the operating point constant in the face of a change in temperature or a change in the β of a transistor. Now we investigate the effects of ac negative feedback. It will be found that negative feedback always serves to reduce voltage gain, but this disadvantage must be weighed against the resulting increase in bandwidth, decrease in distortion, possible increase in input resistance and decrease in output resistance, and quite often, a decrease in noise. Since a common-collector amplifier (emitter follower) involves one hundred percent negative feedback, it is considered in this chapter along with a detailed look at the Darlington connection, which is essentially two common-collectors in cascade.

17-1. GENERAL THEORY OF FEEDBACK

Refer to the block diagram of Fig. 17-1, in which the amplifier's output with S-1 open is V_o, and with S-1 closed is V_o'. Let us consider the following train of events.

The amplifier, having a voltage gain A_v, has a network connected between its output terminals and its input terminals, so that if S-1 is closed some of the output of the amplifier can be fed back to the input.

Consider S-1 open. The output of the feedback network will be zero, so the input to the amplifier will be $V_i + 0 = V_i$. Let the output

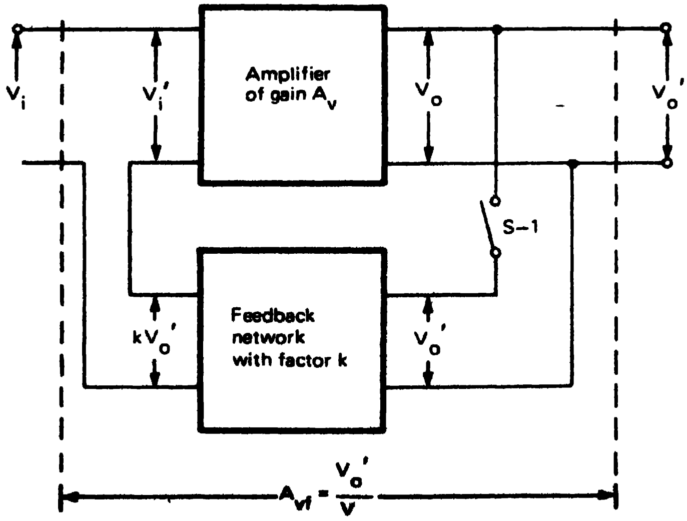


Fig. 17-1. Block diagram of an amplifier with a feedback network that can feed part of the output voltage back in series with the input.

of the amplifier be V_o, so that $A_v = \frac{V_o}{V_i}$. This is the gain of the amplifier without feedback. Similarly, its input resistance is R_i, output resistance R_o, etc.

Now, if S-1 is closed, a certain fraction of the output V_o will be fed back to the input, in series with V_i. This will change the input to the amplifier to some new value, which will in turn change the output from V_o to some new value. How long it will take for the system to settle down to a new equilibrium will depend upon the amplifier and the feedback network. Let the new output of the amplifier (the overall amplifier with feedback) be given by V_o'

The amount of voltage fed back by the feedback network will be kV_o', where k will be less than one (unity) when the feedback network is just made up of passive elements such as resistors, capacitors, inductors. If the network contains an active device, then k can be greater than one.

The new input to the amplifier will then be

$$V_i + kV_o'. \tag{17-1}$$

Whether kV_o' adds to or subtracts from V_i will depend upon the phase relationships, or the sign of k.

The new output from the amplifier will be

$$A_v(V_i + kV_o'). \tag{17-2}$$

But this new output we have designated as V_o'.

$$\therefore V_o' = A_v(V_i + kV_o')$$

$$= A_vV_i + kA_vV_o'.$$

$$\therefore V_o' - kA_vV_o' = A_vV_i$$

$$V_o'(1 - kA_v) = A_vV_i.$$

$$\therefore \frac{V_o'}{V_i} = \frac{A_v}{1 - kA_v}. \tag{17-3}$$

But $\frac{V_o'}{V_i}$ is the gain of the amplifier *with* feedback. Let this be designated as A_{vf}.

$$\text{Then, } A_{vf} = \frac{A_v}{1 - kA_v}. \tag{17-4}$$

Negative Feedback

If k is a negative quantity, the feedback is negative, since the voltage fed back subtracts from the input, yielding a lower output. Hence negative feedback reduces voltage gain.

As an example, say an amplifier has a voltage gain without feedback of $A_v = 40$, and 10% of the output is fed back so as to reduce the input. Then $k = -0.1$ and the gain with feedback is

$$A_{vf} = \frac{A_v}{1 - kA_v} = \frac{40}{1 - (-0.1) \times 40} = \frac{40}{1 + 4} = 8.$$

This is a reduction in gain of 32 or $\frac{32}{40} \times 100\% = 80\%$. Thus, a 10% feedback causes a reduction in gain of 80% in this case. However consider an amplifier with a gain of 400 and a k of -0.1.

$$A_{vf} = \frac{400}{1 - (-0.1) \times 400} = \frac{400}{1 + 40} = 9.76.$$

This is a reduction in gain of 390.24 or $\frac{390}{400} \times 100\% = 97.5\%$. Here, a 10% feedback has caused a reduction in gain of 97.5%. This shows a nonlinear relationship. A certain percentage of feedback does not yield a fixed percentage of reduction in gain. It depends upon the initial value of gain.

Positive Feedback

If the phase of the voltage fed back is such as to increase the input, then k is positive, and the result is positive feedback.

Say $A_v = 40$ and $k = +0.01$. (1% feedback)

$$A_{vf} = \frac{A_v}{1 - kA_v} = \frac{40}{1 - 0.01 \times 40} = \frac{40}{1 - 0.4}$$

$$= \frac{40}{0.6} = 66.6.$$

Since the voltage gain has been increased, the question may arise why use negative feedback?

17-2. REASONS FOR NEGATIVE FEEDBACK

Increased Stability

Consider the equation

$$A_{vf} = \frac{A_v}{1 - kA_v}.$$

If k is negative, and $kA_v >> 1$, then

$$|A_{vf}| \approx \frac{A_v}{kA_v} \approx \frac{1}{k}. \qquad (17\text{-}5)$$

Thus, the gain of the amplifier will be dependent only upon the feedback network, and changes in parameters, such as the beta of a transistor, will affect the gain of the amplifier without feedback, A_v, but will not affect the overall amplifier with feedback. Thus, the gain of the amplifier with feedback has been stabilized against such problems as aging of a tube, or a transistor being replaced by a transistor of a far different value of beta, or temperature variation. Obviously, more stages will have to be added to give the required gain, but this is the price paid for the added stability and other advantages described below. It should be noted that this stability will only come about if the feedback network is made up of stable elements such as resistors, capacitors, and inductors, which is the usual case. If the feedback network has an active device, the advantages will not all materialize. The graph of Fig. 17-2 indicates how the gain of an amplifier is less dependent upon the particular transistor being used when negative feedback is in effect.

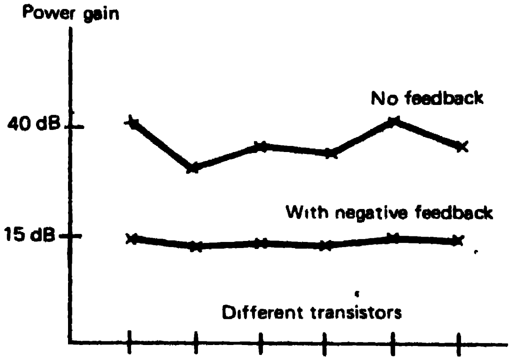


Fig. 17-2. Showing how the power gain using different transistors varies much less with negative feedback than without feedback.

Increased Bandwidth

It can be shown that, if the bandwidth of an amplifier without feedback is given as BW, then the introduction of negative feedback will widen the bandwidth to BW_f.

$$BW_f = BW(1 - kA_v), \qquad (17\text{-}6)$$

where A_v is the midband gain without feedback.

For example, say an amplifier has a BW = 20 kHz, and a gain of $A_v = 40$. Consider the effect of a 1% negative feedback ($k = -0.01$) on the bandwidth and voltage gain.

$$A_{vf} = \frac{A_v}{1 - kA_v} = \frac{40}{1 + 0.01 \times 40} = \frac{40}{1.4} = 28.6.$$

$$BW_f = BW(1 - kA_v) = 20\text{ kHz}(1 + 0.01 \times 40)$$

$$= 20 \times 1.4 = 28\text{ kHz}.$$

Actually, f_2 is increased, and f_1 is decreased, as shown in Fig. 17-3.

It should be noted that the gain is *reduced* by the same factor, $(1 - kA_v)$, that the

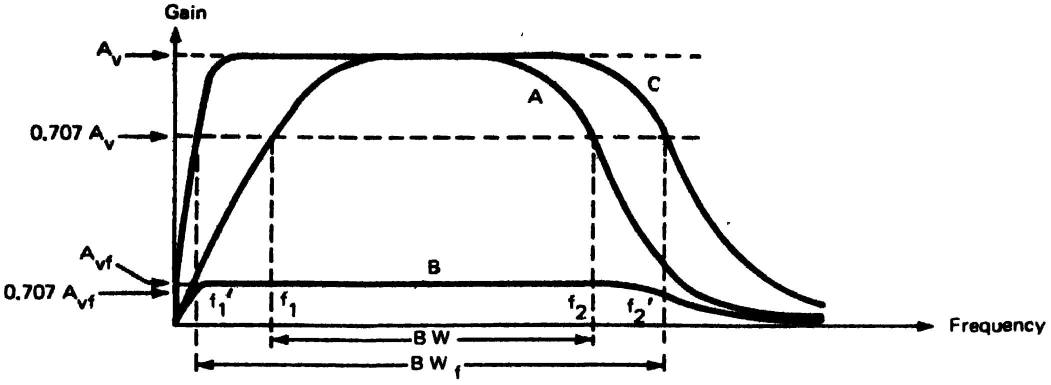


Fig. 17-3. Showing how negative feedback increases bandwidth but reduces midband gain. Curve A is for the amplifier with no feedback. Curve B is with negative feedback but no increase of input signal. Curve C is for negative feedback and increased input signal.

bandwidth is *increased.* Hence, the product of gain and bandwidth *without* feedback is 40 × 20 kHz = 800 kHz, and *with* feedback it is 28.6 × 28 kHz = 800 kHz. This substantiates our earlier findings in Chapter 16 regarding the sacrificing of gain for increased bandwidth.

A physical feeling for the effect of negative feedback can be obtained by considering an amplifier whose voltage gain is 40 at the midband frequencies but drops to 40 × 0.707, or 28, at a frequency of 20 kHz (3 dB point). Now add negative feedback, so that the overall amplifier's gain at the midband frequencies is 28. Remember that the input to the amplifier is *reduced* by the subtraction of some fraction k of the output. If at some frequency, say 20 kHz, the output tends to drop off due to the components in the amplifier, then the input to the amplifier will not be reduced *by the same amount.* In fact, the input to the amplifier will tend to increase (less bucking), keeping the output essentially constant. Since it requires some change in the output to provide this compensating effect at the input, there is still some drop off in output voltage but it is at a much higher frequency (depending upon k) that the 3 dB point will occur.

Decreased Distortion

If an amplifier has a voltage gain A_v and a total harmonic distortion D, then the application of negative feedback, -k, will reduce the distortion to

$$D_f = \frac{D}{1 - kA_v}. \tag{17-7}$$

For example, if D = 10%, when $A_v = 40$ then k = -0.01 will reduce the distortion to

$$D_f = \frac{10}{1 + 0.01 \times 40} = \frac{10}{1.4} = 7.14\%,\ A_{vf} = 28.6.$$

If 10% negative feedback is used, the distortion with feedback will be

$$D_f = \frac{10}{1 + 0.1 \times 40} = \frac{10}{5} = 2\%.$$

However, the gain with feedback would be reduced to

$$A_{vf} = \frac{A_v}{1 - kA_v} = \frac{40}{1 + 0.1 \times 40} = \frac{40}{5} = 8.$$

Thus, if reduction of distortion is of primary interest and gain is "cheap," then a large negative feedback factor would be used.

To get some idea of now negative feedback reduces distortion, consider an amplifier without feedback whose output waveform is as shown in Fig. 17-4(a). V_i is a pure sine wave.

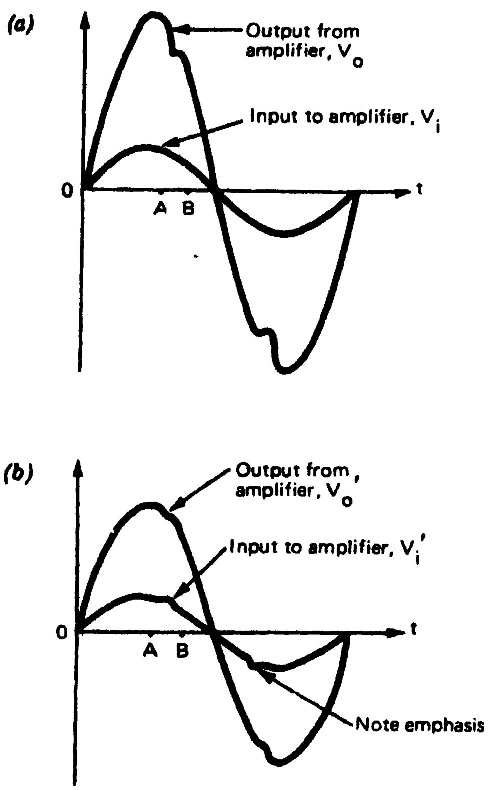


Fig. 17-4. Input and output waveforms for an amplifier. (a) With no feedback, large distortion. (b) With feedback, reduced distortion. Note the input waveform.

If negative feedback is applied, the waveform of V_i' from 0 to A would merely be reduced in size but of the same shape as V_i. However, when the "dip" tries to occur during the interval from A to B, the amount of voltage fed back from the output will tend to *decrease* and reduce the "bucking" effect on the normal input. Thus the input V_i' is effectively *increased* over what it would be without the distortion in the output. The resultant emphasis at the input (see Fig. 17-4(b)) produces a greater output, tending to "fill in the hole" and reduce distortion. Alternatively, the output distortion may be thought of as primarily second harmonic. Then, when the second harmonic goes negative, a portion of this voltage actually *adds* to the input, helping to increase the input and compensate for the loss in output.

Decreased Noise

There are many sources of noise in an amplifier, depending upon whether a tube or transistor is used. The noise, N, can be reduced by a factor of $\frac{1}{1 - kA_v}$ in a similar manner to nonlinear distortion, so that the noise with feedback is given by

$$N_f = \frac{N}{1 - kA_v}. \qquad (17\text{-}8)$$

However, if it is necessary to increase the gain to its original level by the addition of another stage, it is quite possible the overall system will be noisier than it was to start with. If the increase in gain can be accomplished merely by adjustment of circuit parameters, a definite reduction in noise will result from the use of negative feedback.

Input Resistance

It is generally accepted that an increase in input resistance in a voltage amplifier is considered a definite advantage, since it reduces loading on the previous stage and increases the open-circuit voltage gain.

Whereas all the previous factors were improved due to negative feedback, the effect of negative feedback on input resistance (and output resistance) depends upon *the way* in which the output is fed back to the input.

If the output voltage (or current) is fed back in *series* with the input then the input current decreases and the effective input resistance *increases*:

$$R_{if} = R_i(1 - kA_v), \text{ k negative.} \qquad (17\text{-}9)$$

This statement may be verified by referring to the block diagram in Fig. 17-5. R_i is the input resistance with no feedback ($kV_o' = 0$), and R_{if} is the input resistance with feedback. If k is negative, we have negative feedback, so that V_i' is less than V_i, and I_i is reduced.

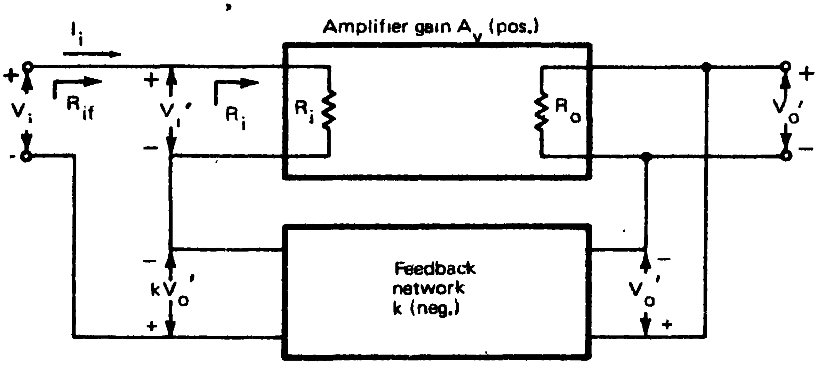


Fig. 17-5. Block diagram of an amplifier with voltage-series feedback that increases input resistance.

$$R_{if} = \frac{V_i}{I_i}, \qquad (17\text{-}10)$$

$$\text{but, } I_i = \frac{V_i'}{R_i}, \text{ and } V_i' = V_i + kV_o'. \qquad (17\text{-}11)$$

Substitution of Eq. 17-11 into Eq. 17-10 yields

$$R_{if} = \frac{R_i}{1 + kA_{vf}}, \qquad (17\text{-}12)$$

$$\text{where } A_{vf} = \frac{V_o'}{V_i} = \frac{A_v}{1 - kA_v}. \qquad (17\text{-}13)$$

Substituting Eq. 17-13 into 17-12 gives the required result of Eq. 17-9:

$$R_{if} = R_i(1 - kA_v).$$

If the output signal (current or voltage) is fed back to the input in *parallel*, the input resistance *decreases*, by an amount depending upon the circuit.

Output Resistance

As with the input resistance, the effect of negative feedback on output resistance is dependent upon the *way* the output is fed back, specifically upon whether output current or output voltage is fed back. If output *voltage* is returned to the input (either in series or shunt), the output resistance decreases:

$$R_{of} = \frac{R_o}{1 - kA_v}, \qquad (17\text{-}14)$$

whereas, if a signal proportional to output *current* is used (in series or shunt), the output resistance increases.

Table 17-1 summarizes the effects of negative feedback on amplifier characteristics

TABLE 17-1 EFFECT OF NEGATIVE FEEDBACK ON AMPLIFIER CHARACTERISTICS

CHARACTERISTIC	TYPE OF FEEDBACK			
	CURRENT-SERIES	VOLTAGE-SERIES	VOLTAGE-SHUNT	CURRENT-SHUNT
Voltage gain	Decreases	Decreases	Decreases	Decreases
Bandwidth	Increases	Increases	Increases	Increases
Harmonic distortion	Decreases	Decreases	Decreases	Decreases
Noise*	Decreases	Decreases	Decreases	Decreases
Input resistance	Increases	Increases	Decreases	Decreases
Output resistance	Increases	Decreases	Decreases	Increases

*Decrease is subject to the condition in Section 17-2.

according to the way in which feedback is accomplished.

Example 17-1. An amplifier with a 1-K ohm input resistance and a 50-K ohm output resistance has a voltage gain of 40. The amplifier is now modified to provide a 10% negative voltage feedback in series with the input. Calculate:

(a) The voltage gain with feedback.
(b) The input resistance with feedback, first using Eq. 17-9, then Eq. 17-12.
(c) The output resistance with feedback.

Solution

(a) $A_{vf} = \dfrac{A_v}{1 - kA_v} = \dfrac{40}{1 - (-0.1)40} = \dfrac{40}{5} = 8.$

(b) $R_{if} = R_i(1 - kA_v) = 1\text{ K }(1 + 4) = 5\text{ K}$ ohms, or

$R_{if} = \dfrac{R_i}{1 + kA_{vf}} = \dfrac{1\text{ K}}{1 + (-0.1)8} = \dfrac{1\text{ K}}{0.2} = 5\text{ K}$ ohms.

(c) $R_{of} = \dfrac{R_o}{1 - kA_v} = \dfrac{50\text{ K}}{1 + 4} = 10\text{ K}$ ohms.

17-3. LOOP GAIN

It should have been noticed that all the calculations above have involved the product kA_v. A special name called the *loop gain* is used to describe this quantity, since it is the combined gain that a voltage at the input is subjected to as it works its way back through the feedback loop to the input. In the above example, the loop gain is −4.

The amount of feedback introduced into an amplifier may be expressed in decibels according to the following definition:

$$F = \text{feedback in dB} = 20 \log \frac{A_{vf}}{A_v}$$

$$= 20 \log \frac{1}{1 - kA_v}, \qquad (17\text{-}15)$$

where k is negative for negative feedback. In Example 17-1,

$$F = 20 \log \frac{8}{40} = -20 \log 5 = -14\text{ dB}.$$

17-4. TYPES OF NEGATIVE FEEDBACK IN TRANSISTOR CIRCUITS

Upon examining Table 17-1, we see that there are four different combinations in which negative feedback may be accomplished. The most desirable from an input-output resistance consideration is the voltage-series type of feedback, but we shall find that it suffers the highest decrease in voltage gain of all the connections. On the other hand, current-shunt feedback has the least desirable effects since it decreases input resistance and increases output resistance. This connection will not be considered here at all.

Current-Series Feedback

In current-series feedback, a voltage is developed which is proportional to the output current. This is called *current* feedback even though it is a voltage that subtracts from the input. (Strictly speaking, current-series feedback stabilizes the transconductance, $G_m = \frac{I_c}{V_i}$, with $k = R_e$. Our discussion makes it appear that the circuit of Fig. 17-6 provides voltage-series feedback, but the equations developed for A_{vf} and R_{if} are valid with $k = \frac{R_e}{R_L}$.)

One of the most common methods of introducing current feedback is to place a resistor in the emitter lead of a CE amplifier. A CE is most often used with negative feedback, since it has such a high gain to start with and can afford to lose some. See Fig. 17-6(a).

When R_e is bypassed with C_e, the output is V_o and the gain without feedback is A_v. R_e provides dc bias stabilization, but no ac feedback. When C_e is removed, an ac voltage will be developed across R_e due to the emitter current (approximately equal to the output current). This voltage will serve to reduce the input voltage between base and emitter, so the output voltage will drop to V_o'. The gain with negative feedback is now A_{vf}.

Approximate Feedback Factor, k

An approximation to the amount of feedback may be obtained by referring to the simplified diagram in Fig. 17-6(b).

$$V_o' = -I_c R_L, \tag{17-16}$$

$$\text{and } V_e = I_e R_e \approx I_c R_e. \tag{17-17}$$

$$\therefore \quad V_e \approx -\frac{R_e}{R_L} V_o'. \tag{17-18}$$

Thus, the fraction of the output voltage fed back in series with the input is given by

$$k \approx \frac{R_e}{R_L}. \tag{17-19}$$

Input Resistance, R_{if}

Equations for R_{if} and A_{vf} for the CE amplifier in Fig. 17-6(a), with an unbypassed emitter resistor, may be obtained using the approximate hybrid equivalent circuit of Section 14-10. However, it is more instructive in our present context to use the circuit of Fig. 17-6(b), since a more accurate expression for k may be found.

$$R_{if} = \frac{V_i}{I_i} = \frac{V_i' + V_e}{I_i} = \frac{V_i'}{I_i} + \frac{V_e}{I_i}.$$

$$\text{But } \frac{V_i'}{I_i} = R_i.$$

Also, $V_e = I_e R_e$, and $I_i = I_e - I_c$.

$$\therefore \ R_{if} = R_i + \frac{I_e R_e}{I_e - I_c} = R_i + \frac{R_e}{1 - \frac{I_c}{I_e}}.$$

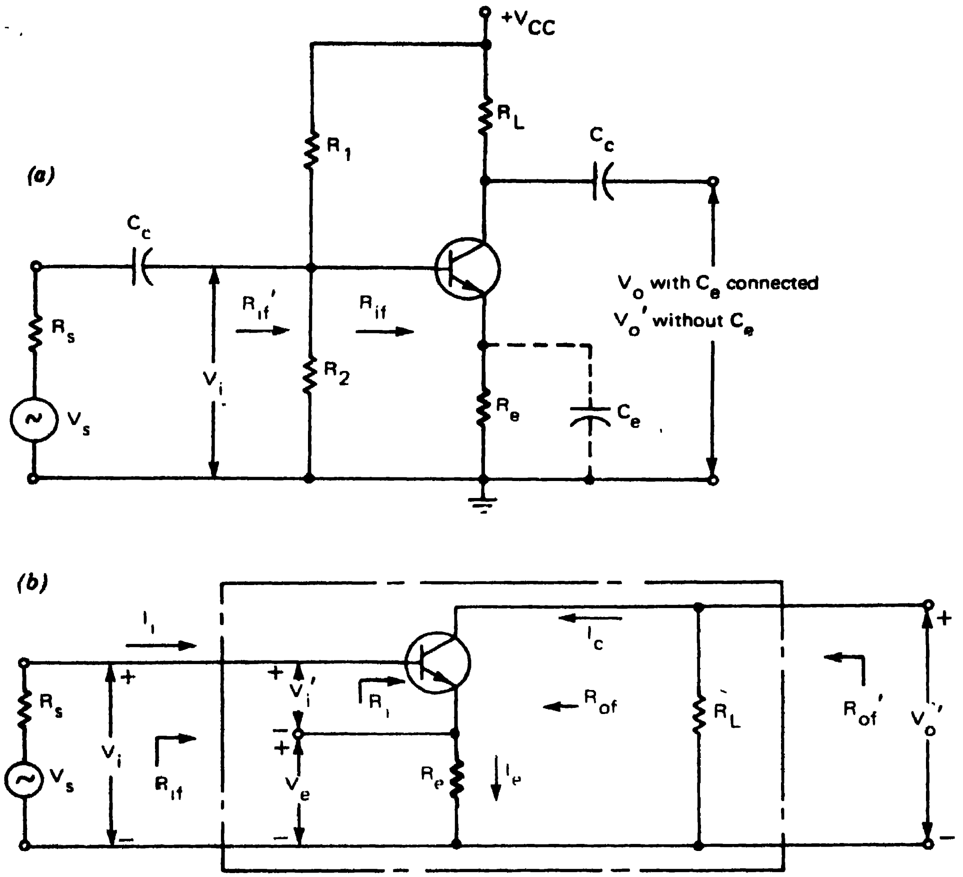

(a)
+V_CC
R_L
C_c
R_1
C_c
R_s
V_s
V_i
R_if'
R_if
R_2
R_e
C_e
V_o with C_e connected
V_o' without C_e
(b)
I_i
R_s
V_s
V_i
R_if
V_i'
V_e
R_i
I_c
R_of
R_L
R_of'
V_o'
R_e
I_e
+
−

Fig. 17-6. (a) CE amplifier with resistor in emitter lead to introduce negative feedback when C_e is removed. (b) Simplified diagram for ac quantities.

But $\frac{I_c}{I_e} = h_{fb} = \frac{h_{fe}}{1 + h_{fe}}$ (for the current directions in Fig. 17-6(b).

$$\therefore R_{if} = R_i + (1 + h_{fe})R_e. \tag{17-20}$$

But $R_i = h_{ie}$ between base and emitter.

$$\therefore R_{if} = h_{ie} + (1 + h_{fe})R_e. \tag{17-21}$$

If the bias resistors $R = R_1 \| R_2$ are taken into account, $R_{if}' = R_{if} \| R$.

Voltage Gain, A_{vf}

Since $A_{vf} = \frac{A_i R_L}{R_{if}}$ (same form as equations in Table 14-4),

and $A_i \approx -h_{fe}$, we obtain

$$A_{vf} = \frac{-h_{fe} R_L}{h_{ie} + (1 + h_{fe})R_e}. \tag{17-22}$$

This compares with $A_v = \dfrac{-h_{fe}R_L}{h_{ie}}$. (17-23)

Evidently, there is a great reduction in gain with negative feedback, and a large increase in input resistance.

Feedback Factor, k

In Section 17-2, we showed that, when a signal is fed back in series with the input, the input resistance is given by Eq. 17-9:

$$R_{if} = R_i(1 - kA_v).$$

But, from Eq. 17-20, we have

$$R_{if} = R_i + (1 + h_{fe})R_e.$$

Equating these two expressions and using $A_v = \dfrac{-h_{fe}R_L}{R_i}$, we can solve for k:

$$k = \frac{1 + h_{fe}}{h_{fe}} \times \frac{R_e}{R_L} = \frac{R_e}{h_{fb}R_L}. \qquad (17\text{-}24)$$

Since $h_{fb} \approx 1$, $k \approx \dfrac{R_e}{R_L}$, as found in Eq. 17-19.

That the feedback factor is somewhat larger than $\dfrac{R_e}{R_L}$ is not surprising, since the emitter current is slightly larger than the output collector current.

Output Resistance, R_{of}

The expression for the output resistance, R_{of}, looking back into the collector is involved and contains R_s, R_e, and all the h-parameters. For values of R_e in the order of R_s and h_{ie}, an approximate expression for R_{of} is

$$R_{of} \approx \frac{1 + h_{fe}}{h_{oe}} \approx \frac{1}{h_{ob}}. \qquad (17\text{-}25)$$

This is usually a very large value, in the megohms range, so the overall output resistance, R_{of}', including the load resistor, is approximately R_L.

Example 17-2. For the amplifier in Fig. 17-6, the transistor has $h_{fe} = 50$ and $h_{ie} = 1$ K ohm. With $R_L = 1$ K ohm, $R_e = 100$ ohms, $R = R_1 \| R_2 = 10$ K ohms, calculate the following:

(a) The gain A_v if R_e were adequately bypassed, and R_i.
(b) The approximate feedback factor k (C_e removed).
(c) The input resistance with feedback, R_{if}, and R_{if}' using Eq. 17-21.
(d) R_{if} using Eq. 17-9, accounting for the difference from (c).
(e) A_{vf}.
(f) The loop gain and the feedback in decibels.

Solution

(a) $A_v = \dfrac{A_i R_L}{R_i} \approx \dfrac{-h_{fe}R_L}{h_{ie}} = \dfrac{-50 \times 1\text{ K}}{1\text{ K}} = -50.$

$R_i \approx h_{ie} = 1$ K ohm.

(b) $k \approx \dfrac{R_e}{R_L} = \dfrac{100}{1000} = 0.1$ or 10%.

(c)
$$\begin{aligned} R_{if} &= h_{ie} + (1 + h_{fe})R_e \\ &= 1\text{ K} + (1 + 50)100 \\ &= 6.1\text{ K ohms.} \\ R_{if}' &= R_{if} \| R \\ &= 6.1\text{ K} \| 10\text{ K} = 3.8\text{ K ohms.} \end{aligned}$$

(d)
$$\begin{aligned} R_{if} &= R_i(1 - kA_v) = 1\text{ K}\,[1 - (0.1)(-50)] \\ &= 1\text{ K} \times 6 = 6\text{ K ohms.} \end{aligned}$$

The difference from (c) arises from the fact that we used $k = \dfrac{R_e}{R_L} = 0.1$.

In fact, $k = \dfrac{1 + h_{fe}}{h_{fe}} \times \dfrac{R_e}{R_L} = \dfrac{51}{50} \times 0.1 = 0.102.$

Then, $R_{if} = R_i(1 - kA_v) = 1\text{ K}(1 + 5.1) = 6.1$ K ohms.

(e) $A_{vf} = \dfrac{-h_{fe}R_L}{R_{if}} = \dfrac{-50 \times 1\text{ K}}{6.1\text{ K}} = -8.2.$

(f) Loop gain $= kA_v = 0.102 \times (-50) = -5.1.$

$$F = 20 \log \left(\frac{A_{vf}}{A_v}\right) = 20 \log \frac{-8.2}{-50} = -20 \log 6.1$$
$$= -15.7 \text{ dB}.$$

Note: In this section, k was positive, but yet we had negative feedback. This is because of the 180° phase shift in the CE making A_v negative, so that kA_v is a negative quantity.

If it is found that a certain value of R_e is required for bias stabilization but not all of R_e is required for negative ac feedback, just part of R_e can be bypassed as shown in Fig. 17-7. Then the feedback factor is approximately equal to $\frac{R_{e1}}{R_L}$.

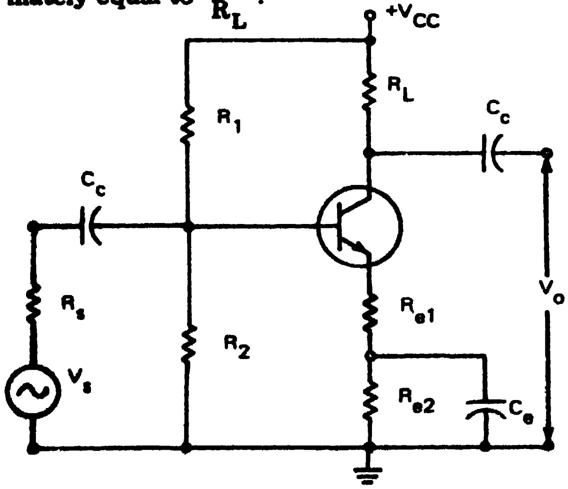


Fig. 17-7. Amplifier with emitter bias due to R_{e1} + R_{e2} but negative ac feedback factor $k \approx R_{e1}/R_L$.

Voltage-Series Feedback

A special case of the CE amplifier with the emitter unbypassed is the common-collector amplifier. This is shown in Fig. 17-8, where the output voltage is taken across $R_e = R_L$, and the negative feedback factor is

$$k = -\frac{R_e}{R_L} = -1. \tag{17-26}$$

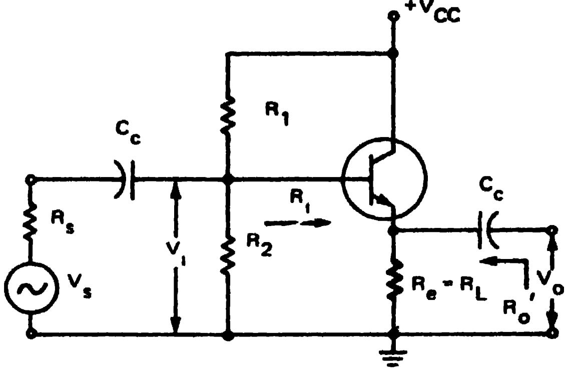


Fig. 17-8. A common-collector amplifier or emitter-follower that exhibits 100% negative feedback.

Thus, the common-collector amplifier or emitter-follower (since the voltage at the emitter follows the input voltage) exhibits 100 percent negative feedback. This is voltage-series feedback, since the voltage developed in the output is in series with the input voltage as far as the base-emitter-junction is concerned.

Common-Collector Characteristics

$$A_i = \frac{I_e}{I_b} = \frac{I_b + I_c}{I_b} = \frac{I_b + h_{fe}I_b}{I_b} = 1 + h_{fe}. \tag{17-27}$$

$$R_i = h_{ie} + (1 + h_{fe})R_e = h_{ie} + (1 + h_{fe})R_L. \tag{17-28}$$

$$A_v = \frac{A_i R_L}{R_i} = \frac{(1 + h_{fe})R_L}{R_i}$$
$$= \frac{(1 + h_{fe})R_L}{h_{ie} + (1 + h_{fe})R_L} = 1 - \frac{h_{ie}}{R_i}. \tag{17-29}$$

$$R_o = \frac{h_{ie} + R_s}{1 + h_{fe}}, \tag{17-30}$$

and $R_o' = R_o \| R_L$. (17-31)

Note: A resistor may be inserted in the collector lead of the emitter-follower to protect the transistor in the event of a short circuit across the output. This has negligible effect on the above equations.

For the values given in Example 17-2, with $R_s = 600$ ohms and $R_L = 1$ K ohm, we find $A_i = 51$, $R_i = 52$ K ohms, $A_v = 0.98$, $R_o = 31$ ohms, and $R_o' = 30$ ohms.

The low gain due to the high degree of negative feedback ensures the common collector circuit an excellent high-frequency capability. Its low output resistance and high input resistance makes it ideal as the last stage in a cascaded amplifier system.

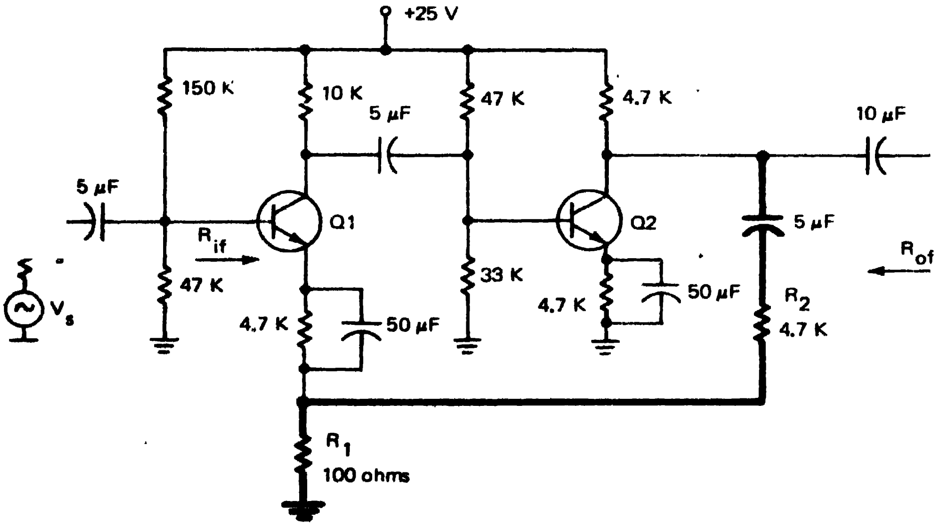


Fig. 17-9. Two-stage amplifier that uses voltage-series feedback.

A voltage-series feedback circuit can involve two transistors as shown in Fig. 17-9.

Here the feedback factor, $k = \dfrac{-R_1}{R_1 + R_2}$

$$\frac{-100}{100 + 4700}$$

$$= \frac{-1}{48}.$$

The overall gain of 4,580 without feedback is thus reduced to 47.5, the input resistance R_{if} increased to 106 K ohms from 1.1 K ohms, and the output resistance decreased from 2.35 K ohms to 25 ohms. ($h_{ie} = 1.1$ K ohms, $h_{fe} = 50$.)

Voltage-Shunt Feedback

A voltage-shunt feedback circuit is illustrated in Fig. 17-10. This is similar to the circuit used in Chapter 12 to provide stabilization of the operating point against temperature variations. However, in that application, to avoid degenerative or negative feedback, R_B was split into two, and a capacitor connected from their junction to ground. Here, a current that is proportional to the output voltage is fed back to the input. Since $V_o >> V_i$, the feedback current $I_f \approx \frac{V_o}{R_B}$, so that the proportionality constant or feedback factor $k \approx \frac{1}{R_B}$.

$$R_{of}' = R_{of} \| R_L. \qquad (17\text{-}37)$$

For h_{ie} = 1 K ohm, h_{fe} = 50, R_L = 1 K ohm, R_B = 40 K ohms, R_s = 600 ohms, we obtain R_L' = 975 ohms, $A_v = -49$, R_{if} = 440 ohms, $A_{vsf} = -21$, R_{of} = 2.1 K ohms, $R_{of}' -$ 680 ohms.

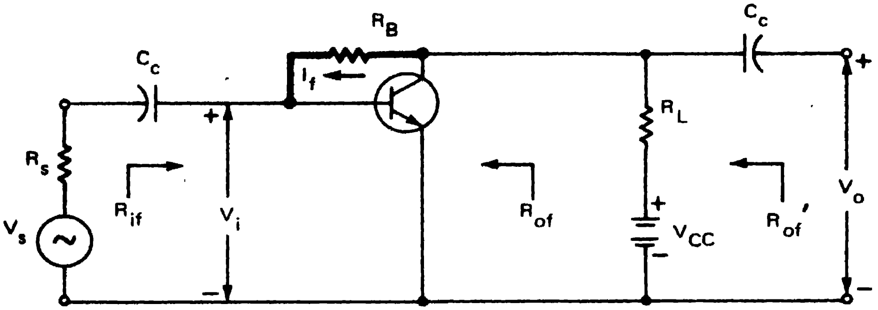


Fig. 17-10. CE amplifier with voltage-shunt feedback provided by R_B.

The reduction in input and output resistance comes from the Miller effect due to R_B. That is,

$$R_{if} = h_{ie} \| \frac{R_B}{1 - A_v}, \qquad (17\text{-}32)$$

$$\text{where } A_v = \frac{V_o}{V_i} = \frac{-h_{fe} R_L'}{h_{ie}}, \qquad (17\text{-}33)$$

$$\text{and } R_L' = R_B \| R_L. \qquad (17\text{-}34)$$

$$A_{vsf} = \frac{V_o}{V_s} = A_v \frac{R_{if}}{R_s + R_{if}}. \qquad (17\text{-}35)$$

$$R_{of} \approx \frac{R_B}{R_s} \times \frac{(R_s + h_{ie})}{h_{fe}}. \qquad (17\text{-}36)$$

17-5. THE DARLINGTON CONNECTION

Let us now investigate a method of increasing the input resistance to a transistor amplifier. This is desirable in many cases, especially when a high gain (50,000 and up) operational amplifier is required which depends upon a very high input resistance for its operation as an integrator or summing amplifier in analog applications.

The common-collector or emitter follower was found to have 100% negative feedback, resulting in the largest input resistance and lowest output resistance of all three amplifiers. The Darlington connection can be considered as two cascaded emitter followers, with the first stage having an infinite emitter resistance, as shown in Fig. 17-11.

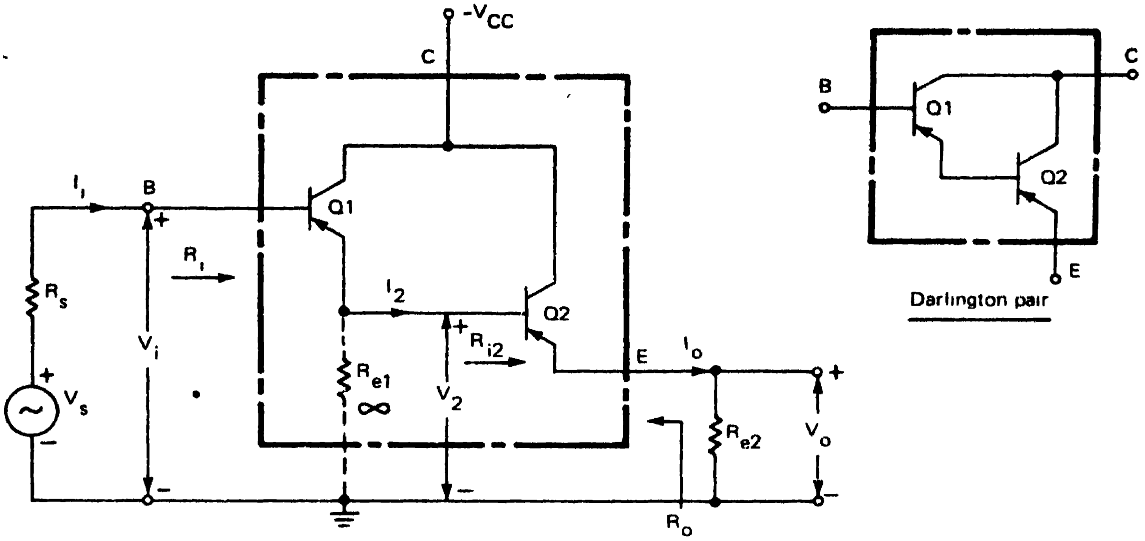


Fig. 17-11. High-input-resistance amplifier utilizing a Darlington pair. Bias resistors have not been included.

Note: The biasing network for Q1 is not included, but will be added later. Assume the operating point for it to be correct.

Since there is no connection to the external circuitry from the emitter of Q1 or the base of Q2, these two transistors can be fabricated as a single device and sold as a Darlington pair with three leads. An example of this is the 2N999, which has a voltage capability of $V_{CE} = 60$ V and an h_{fe} between 7000 and 70,000 at a collector current of 100 mA. Power capability is 500 mW.

Analysis of the circuit in Fig. 17-11 proceeds from right to left, with R_{e2} the load for Q2.

The current gain for this conventional emitter follower is given by

$$A_{i2} = \frac{I_o}{I_2} = \frac{1 + h_{fe}}{1 + h_{oe}R_{e2}} \approx 1 + h_{fe}, \qquad (17\text{-}38)$$

since $h_{oe}R_{e2} \leq 0.1$, and the input resistance by

$$R_{i2} = h_{ie} + (1 + h_{fe})R_{e2} \approx (1 + h_{fe})R_{e2}, \qquad (17\text{-}39)$$

since h_{ie} is usually much smaller than $h_{fe}R_{e2}$.

But R_{i2} is the effective load resistance (emitter resistance) for the first stage Q1, so that the current gain for the first stage is given by

$$A_{i1} = \frac{I_2}{I_i} = \frac{1 + h_{fe}}{1 + h_{oe}R_{i2}} = \frac{1 + h_{fe}}{1 + h_{oe}(1 + h_{fe})R_{e2}}$$

$$\approx \frac{1 + h_{fe}}{1 + h_{oe}\,h_{fe}R_{e2}}, \text{ since } h_{oe}R_{e2} \leq 0.1. \qquad (17\text{-}40)$$

Thus, the overall current gain $A_i = \frac{I_o}{I_i}$

$$= \frac{I_o}{I_2} \times \frac{I_2}{I_i}.$$

$$\therefore A_i = A_{i2} \times A_{i1} \approx (1 + h_{fe})\frac{(1 + h_{fe})}{1 + h_{oe}h_{fe}R_{e2}}$$

$$\therefore A_i \approx \frac{(1+h_{fe})^2}{1+h_{oe}h_{fe}R_{e2}}. \qquad (17\text{-}41)$$

Similarly, the input resistance can be obtained:

$$R_i = \frac{A_i R_{e2}}{A_v} \approx A_i R_{e2} \approx \frac{(1+h_{fe})^2 R_{e2}}{1+h_{oe}h_{fe}R_{e2}}. \qquad (17\text{-}42)$$

The overall voltage gain of the Darlington circuit is less than unity, since it consists of two emitter followers in cascade, each having A_v less than unity:

$$A_v = \frac{V_o}{V_i} \approx 1 - \frac{h_{ie}}{R_{i2}} \qquad (17\text{-}43)$$

$$\text{Also, } R_o \approx \frac{R_s + h_{ie}}{(1+h_{fe})^2} + \frac{h_{ie}}{1+h_{fe}}. \qquad (17\text{-}44)$$

It should be noted that, for convenience, it has been assumed that the h-parameters are the same for each transistor in writing the above equations. In general, this will not be true, since the collector (emitter) current of Q1 is the base current of Q2, so that, unless Q2 is a relatively large transistor (with a large base current requirement), Q1 will be operated at a low level of collector (emitter) current, at which point its h_{fe} may be smaller than that of Q2.

Example 17-3. To get some idea of the values of input resistance etc., let us use the following h-parameters to calculate the properties of the above Darlington amplifier, with $R_{e2} = 3$ K ohms, $R_s = 3$ K ohms, $h_{ie} = 1{,}100$ ohms, $h_{re} = 2.5 \times 10^{-4}$; $h_{fe} = 50$; $h_{oe} = 25 \times 10^{-6}$ mhos. (Note: $h_{oe}R_{e2} = 0.075 < 0.1$.)

Solution

$$A_i \approx \frac{(1+h_{fe})^2}{1+h_{oe}h_{fe}R_{e2}}$$

$$= \frac{(1+50)^2}{1+25\times 10^{-6}\times 50\times 3000} = 548.$$

$$R_i \approx \frac{(1+h_{fe})^2 R_{e2}}{1+h_{oe}h_{fe}R_{e2}} = A_i R_{e2} = 548 \times 3000$$

$$= 1.64 \text{ M ohms.}$$

$$A_v \approx 1 - \frac{h_{ie}}{R_{i2}} = 1 - \frac{1{,}100}{(1+50)3000} = 0.993.$$

$$R_o \approx \frac{R_s + h_{ie}}{(1+h_{fe})^2} + \frac{h_{ie}}{1+h_{fe}} = \frac{3000+1100}{(1+50)^2}$$

$$+ \frac{1100}{1+50} = 23.2 \text{ ohms.}$$

These values should be compared with a single-stage emitter follower having the same values of R_s, R_e and h-parameters:

$A_i = 51$; $R_i = 154$ K ohms; $A_v = 0.993$;

$R_o = 80.5$ ohms.

Thus, the Darlington connection has *increased* the input resistance by a factor of 10, *reduced* the output resistance by a factor of 4, and *increased* the current gain by a factor of 11. The voltage gain is the same. These are all advantages and show the vast improvement over a single-stage emitter follower.

17-6. BIASING THE DARLINGTON AMPLIFIER

It was mentioned that the biasing network had been omitted from Q1 to simplify analysis, but for a practical circuit we must examine what the effect of such a biasing network will have on the desirably large input resistance provided by the Darlington amplifier.

Consider the "standard" voltage divider network for biasing Q1, as shown in Fig. 17-12. Now the input resistance is $R_i' = R_i \,||\, R$, where

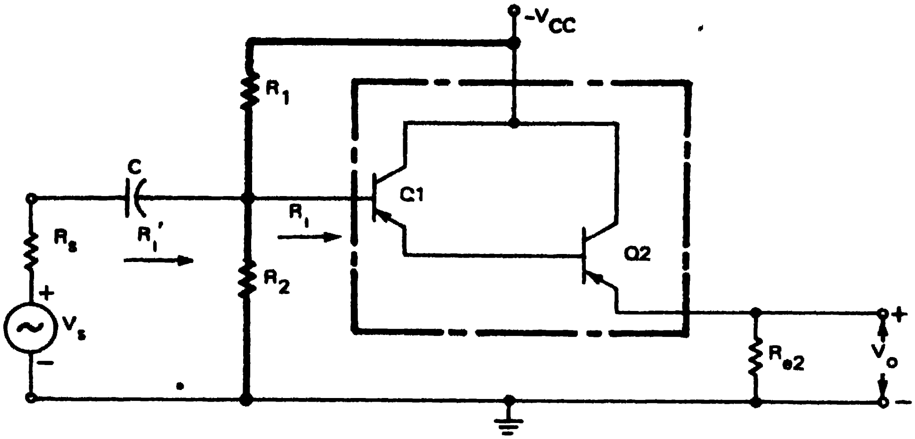


Fig. 17-12. Darlington amplifier with biasing resistors R_1 and R_2 causing a reduction in input resistance from R_i to R_i'

$R = R_1 \| R_2$. But the parallel combination of R_1 and R_2 will be much less than R_i, since the dc bias current into the base of Q1 is normally in the region of 0.1 mA, placing an upper limit of a few hundred kilohms on R. Hence, the effective input resistance R_i' will be appreciably less than R_i, defeating the very purpose of the Darlington connection.

This can be partially overcome by adding R_3 as shown in Fig. 17-13. (Ignore C_B for the present.) Now R is increased to $R_1 \| R_2 + R_3$, but this is still much less than R_i (1.64 M ohms), so another method is required.

The Bootstrapping Principle

Actually, the addition of R_3 in Fig. 17-13 is an

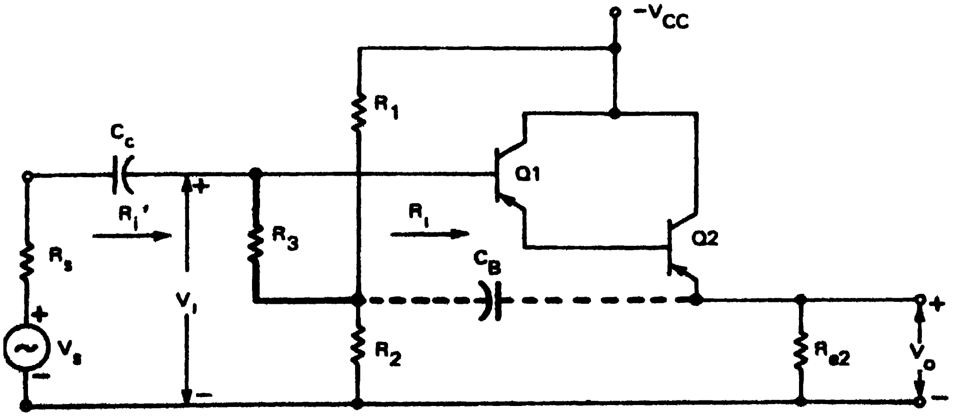


Fig. 17-13. Darlington amplifier with modified bias circuit.

intermediate step to applying the bootstrapping principle. Consider the addition of C_B shown dotted in the above diagram. Its reactance is considered to be negligible at the lowest frequency at which the amplifier is to be used. Hence, C_B directly couples the output voltage V_o to the lower side of R_3, while V_i, which is of the same phase as V_o and substantially equal to V_o, is on the upper side of R_3. There is then effectively no ac voltage drop across R_3, so that R_3 draws negligible ac current from the input V_i. The effective value of R_3 with this coupling arrangement has now been increased, by the Miller effect, to a value given by

$$R_{3,\text{eff}} = \frac{R_3}{1 - A_v}, \text{ where } A_v \text{ is positive.} \quad (17\text{-}45)$$

For example, A_v was calculated to be 0.993 for R_{e2} = 3 K ohms. Say R_3 = 70 K ohms (determined by the quiescent base current requirement), then $R_{3,\text{eff}}$ is

$$R_{3,\text{eff}} = \frac{70\text{ K}}{1 - 0.993} = \frac{70\text{ K}}{0.007} = 10\text{ M ohms.}$$

That is, in an emitter follower where A_v is close to unity, $R_{3,\text{eff}}$ becomes very large and reduces the bias circuit loading on the input signal V_i.

In the above equation for $R_{3,\text{eff}}$, the effect of A_v approaching 1 is termed "bootstrapping." This term is appropriate since, when V_i increases, V_o also increases by practically the same amount. Thus both ends of R_3 increase by the same potential; the effect is as if R_3 were "pulling itself up by its bootstraps." Since no additional voltage drop occurs across R_3, it cannot be considered as being a load on the input.

With the above assumptions of R_3 = 70 K ohms, the input resistance to the amplifier complete with biasing circuit, R_i', should be $R_i \| R_{3,\text{eff}} = 1.64\text{ M} \| 10\text{ M} \approx 1.4\text{ M ohms.}$

Actually, R_i is slightly different, since the effective emitter resistance for Q2 is now given by R_{e2} in parallel with $R_1 \| R_2$. However, R_{e2} will normally be much smaller than $R_1 \| R_2$, so there is not much difference.

It should be mentioned that the high input resistance of the Darlington connection could be obtained using a field-effect transistor (FET) (Chapter 21). However, since an FET has a somewhat poorer high frequency response than a bipolar transistor, applications remain for the Darlington connection.

17-7. SUMMARY

(1) If an amplifier having a voltage gain A_v has some of the output voltage fed back in series with the input, the voltage gain with feedback is $A_{vf} = \frac{A_v}{1 - kA_v}$, where k is the feedback factor, negative for negative feedback, and positive for positive feedback.

(2) Although negative feedback reduces voltage gain, the result is a stabilization of gain, since $A_{vf} \approx \frac{1}{k}$.

(3) Advantages of negative feedback include extended bandwidth, reduced harmonic distortion, and possibly a reduction in noise.

(4) The effect of negative feedback on input and output resistance depends upon the type of negative feedback, as summarized in Table 17-1. The four types are current- and voltage-series, and current- and voltage-shunt.

(5) Loop gain is the product kA_v and is negative for negative feedback.

(6) Feedback in dB = $20 \log \frac{A_{vf}}{A_v} = 20 \log \frac{1}{1 - kA_v}$.

(7) An unbypassed resistor in the emitter lead of a CE amplifier is an example of current-series

feedback that increases input resistance, but also increases output resistance. The feedback factor $k \approx R_e/R_L$.

(8) The common-collector or emitter follower is an example of voltage-series feedback. It inherently uses 100% negative feedback and is the one connection that simultaneously increases input resistance and decreases output resistance. Its voltage gain is a little less than unity.

(9) Voltage-shunt feedback, such as a resistor connected from collector-to-base in a CE amplifier, reduces both input and output resistance.

(10) A Darlington connection involves tying together the collectors of two transistors and the emitter of the first to the base of the second to provide a three-terminal device with very high current gain.

(11) The input resistance and current gain of a Darlington amplifier are very high, its output resistance very low, and its voltage gain approximately one.

(12) To maintain high input resistance, the biasing of a Darlington amplifier must involve "bootstrapping," in which practically equal voltages exist at the ends of a resistor, so that it offers practically no load.

QUESTIONS

Q17-1. (a) What is the general equation for the voltage gain of an amplifier with feedback?
(b) Is A_{vf} necessarily smaller than A_v? Explain.

Q17-2. (a) Why is the voltage gain of an amplifier with negative feedback smaller than with no feedback?
(b) Does a certain percentage of feedback produce a fixed percentage of reduction in gain? Why or why not?

Q17-3. List as many advantages of negative feedback as you can think of, indicating any possible restrictions or disadvantages associated with each.

Q17-4. Explain in your own words how negative feedback can increase the value of bandwidth in an amplifier.

Q17-5. (a) What types of negative feedback are there?
(b) Which ones always produce an increase in input resistance?
(c) Which type increases input resistance *and* decreases output resistance?

Q17-6. (a) What is *loop gain*?
(b) Why is negative feedback expressed in decibels a negative number?

Q17-7. (a) How is current-series feedback introduced in a CE amplifier?
(b) What is an approximate expression for the feedback factor?
(c) What happens to the voltage gain, input resistance, and output resistance in a CE amplifier with current-series negative feedback?

Q17-8. How may the resistance used for dc negative feedback (in bias point stabilization) be made different from the resistance used for ac negative feedback in a CE amplifier with current-series feedback?

Q17-9. (a) What type of feedback is the common collector (or emitter follower) an example of?
(b) How much negative feedback is used?
(c) What is the approximate resultant voltage gain?
(d) Why should an emitter follower have an excellent high frequency capability? (Include the Miller effect.)

Q17-10. (a) What phase relationship exists between V_o and V_s at the mid-frequency range in Fig. 17-9?

(b) What is the phase relationship between the voltage across R_1 and V_s?
(c) Explain how R_1 introduces negative feedback.

Q17-11. Even though the circuit of Fig. 17-10 reduces input resistance and voltage gain what advantage does the circuit have over conventional biasing?

Q17-12. (a) What does the Darlington connection consist of essentially?
(b) What are the primary features of a Darlington amplifier?

Q17-13. (a) Explain in your own words what is meant by "bootstrapping."
(b) Why is it found desirable to use bootstrapping in the bias circuit of a Darlington amplifier?

PROBLEMS

P17-1. An amplifier has a voltage gain of 100. A technician decides that 10% negative feedback should be employed to reduce distortion.
(a) What will be the voltage gain with feedback?
(b) Finding the voltage gain to be excessively low with 10% negative feedback, the technician believes a doubling of voltage gain will result if he reduces the negative feedback to 5%. By how much does the actual voltage gain differ from the expected value?
(c) What should he have reduced the feedback to in order to double the gain that existed in (a)?

P17-2. Repeat P17-1 with an initial voltage gain of 50.

P17-3. (a) If an amplifier has a bandwidth of 200 kHz and a voltage gain of 100, what will be the new bandwidth and gain if 5% negative feedback is introduced?
(b) What is the product of gain and bandwidth before and after adding negative feedback in (a)?
(c) What should be the amount of feedback if the bandwidth is to be restricted to 1 MHz?

P17-4. Repeat P17-3 with an initial voltage gain of 50.

P17-5. An amplifier's total harmonic distortion is reduced from 8% to 2% when 5% negative feedback is used.
(a) What was the initial voltage gain of the amplifier when the distortion was 8%?
(b) What was the gain with 2% distortion?
(c) How much distortion would there be if the gain with feedback were reduced to 10?

P17-6. Negative feedback is to be used to reduce the noise from an amplifier by 90%.
(a) What must be the percentage negative feedback to accomplish this if the initial voltage gain is 50?
(b) What will be the voltage gain with feedback?

P17-7. An amplifier with a 2.2-K ohm input resistance and 40-K ohm output resistance has a voltage gain of 80. The amplifier is now modified to provide 10% negative voltage feedback in series with the input. Calculate:
(a) The voltage gain with feedback.
(b) The input resistance.
(c) The output resistance.

If the feedback were effected using current-series feedback, which of the above (if any) would change?

P17-8. An amplifier has -20 dB of feedback introduced.
(a) What is the ratio of voltage gain before feedback to the gain with feedback?
(b) Determine the loop gain.
(c) If k = -5%, what is the gain before and after feedback?

P17-9. Repeat P17-8 with –40 dB of feedback.

P17-10. An amplifier with a voltage gain of 180 has +20 dB of feedback introduced. Calculate:

(a) The loop gain.
(b) The percentage and type of feedback.
(c) The voltage gain with feedback.

P17-11. An amplifier with current-series feedback has the form shown in Fig. 17-6(a). The transistor is 2N217 with h_{fe} = 100, h_{ie} = 2 K ohms. R_L = 470 ohms, and R_e consists of two 100-ohm resistors, so that both or just one may be bypassed by C_e, which has negligible reactance. R_2 = 5.6 K ohms and R_1 = 15 K ohms.

(a) Calculate the voltage gain, R_i, and R_i'. with R_e completely bypassed with C_e. (Use approximate equations.)
(b) Assume C_e bypasses one 100-ohm resistor. Determine the approximate feedback factor k.
(c) Calculate the input resistance with feedback, R_{if}, using first Eq. 17-21 then 17-9. Why any difference?
(d) Determine R_{if}'.
(e) Calculate A_{vf}, the loop gain, and feedback in decibels.

P17-12. Repeat P17-11,(b) through (e), assuming R_e = 200 ohms.

P17-13. A transistor with h_{ie} = 1.5 K ohms and h_{fe} = 75 is used in an emitter-follower circuit, as in Fig. 17-8. If R_e = R_L = 860 ohms and $R_1 \| R_2$ = 20 K ohms, calculate:

(a) A_i, (b) R_i, (c) R_i',
(d) A_v, (e) R_o, (f) R_o', assuming R_s = 1 K ohm.

P17-14. (a) For the data given in P17-13, determine the value of R_L to provide an input resistance of 100 K ohms.

(b) What is the voltage gain for this value of R_L?
(c) What will be the new input resistance and voltage gain if the amplifier drives a 1-K ohm load?

P17-15. An amplifier employs collector-to-base biasing to provide bias point stabilization. No bypass capacitor is used, so that negative voltage-shunt feedback results, as in Fig. 17-10. R_B = 80 K ohms, R_L = 1 K ohm, h_{ie} = 2 K ohms, and h_{fe} = 75. If the source resistance is 1 K ohm, calculate:

(a) R_{if}, (b) A_v, (c) R_{of},
(d) R_{of}'.

P17-16. A Darlington amplifier as in Fig. 17-12 employs two transistors whose parameters are assumed to be equal; h_{fe} = 100, h_{ie} = 2 K ohms, h_{oe} = 50 × 10^{-6} mho. If R_{e2} = 2 K ohms, $R_1 \| R_2$ = 300 K ohms and R_s = 1 K ohm, calculate:

(a) A_i, (b) R_i, (c) R_i'.
(d) A_v, (e) R_o.

P17-17. More realistic values for the problem in P17-16 would take into account the fact that the transistors will operate at different values of collector current. Repeat P17-16 using:

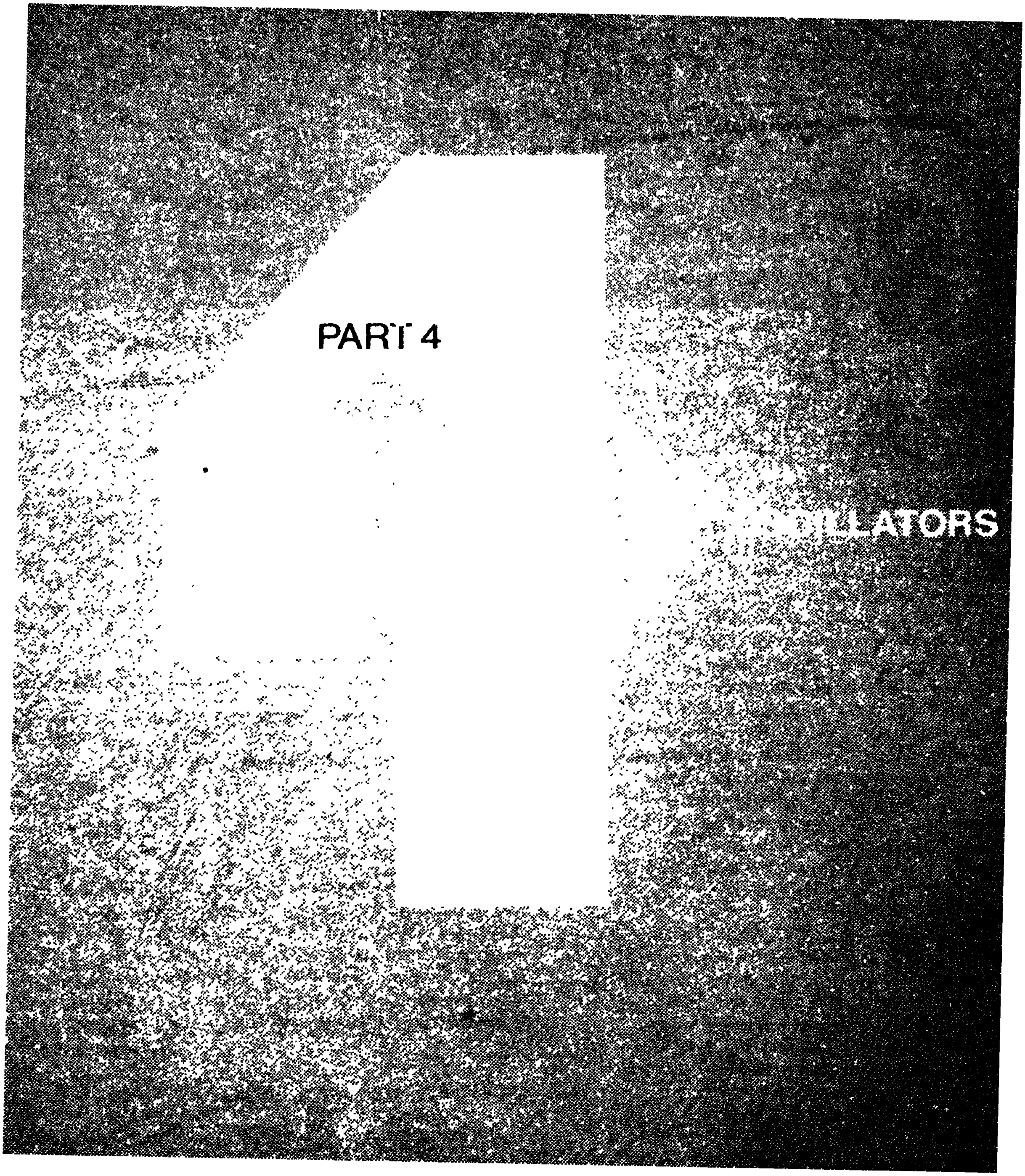
h_{fe1} = 80, h_{fe2} = 100, h_{ie1} = 6 K ohms,

h_{ie2} = 1 K ohm, h_{oe1} = 30 × 10^{-6} mho,

h_{oe2} = 50 × 10^{-6} mho.

(Examine the development of the equations for A_i, etc. to see what changes to make in the subscripts.)

P17-18. The bias circuit in P17-16 is modified to that shown in Fig. 17-13, with R_3 = 40 K ohms. Determine the new value of input resistance, R_i'.



PART 4

LLATORS

18

TRANSISTOR OSCILLATORS AND MULTIVIBRATORS

18-0. INTRODUCTION

In the previous chapter, we examined the effect of feeding some of the output of an amplifier back to the input in such a way as to reduce the input. The result was a reduction of overall voltage gain, although numerous beneficial effects were cited, such as increased stability, widened bandwidth, reduced distortion, and, for some circuits, increased input resistance and decreased output resistance. This was termed negative feedback. We now examine the result of positive feedback and the eventual oscillation that can occur, providing a source of ac voltage at a desired frequency called an oscillator.

Strictly speaking, an oscillator provides a sinusoidal output, at frequencies ranging from the audio band up to several gigahertz (1 GHz = 1000 MHz). In addition, the frequency may be variable. At low frequencies, the oscillators usually involve R-C components for frequency selection, but higher frequencies rely upon LC resonant circuits. Nonsinusoidal oscillators exist and are often called relaxation oscillators. They may involve an active device that has a negative-resistance effect such as the tunnel diode, or a unijunction transistor, a device to be considered in Chapter 28. Other nonsinusoidal oscillators that will be covered in this chapter are multivibrators or square-wave oscillators. Although these are switching-type circuits and are found in pulse or digital circuits, they do involve positive feedback, so their discussion is appropriate here, although we shall again consider them, in integrated circuit form, in Chapter 26.

18-1. EFFECT OF POSITIVE FEEDBACK

In Chapter 17, we examined the block diagram of Fig. 18-1, in which the loop gain, kA_v, was negative.

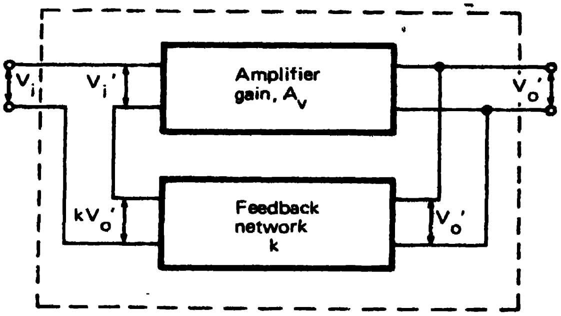


Fig. 18-1. Block diagram of a feedback system.

The overall gain with feedback, V_o'/V_i, was shown to be

$$A_{vf} = \frac{V_o'}{V_i} = \frac{A_v}{1 - kA_v} \qquad (18\text{-}1)$$

where A_v is the amplifier gain, V_o'/V_i', and k is the portion of the output voltage fed back to the input. Obviously, if kA_v is negative, the overall gain is reduced, but what happens if kA_v is positive, to provide positive feedback? The answer depends upon how much feedback is introduced, as shown by the following:

Consider $A_v = 40$ and $k = 0.01, 0.02, 0.025$:

$$A_{vf} = \frac{40}{1 - 0.01 \times 40} = \frac{40}{1 - 0.4} = \frac{40}{0.6}$$

$$= 66.7 \text{ for } k = 0.01.$$

$$\frac{40}{1 - 0.02 \times 40} = \frac{40}{1 - 0.8}$$

$$= \frac{40}{0.2} = 200 \text{ for } k = 0.0$$

$$\frac{40}{1 - 0.025 \times 40} = \frac{40}{1 - 1}$$

$$= \frac{40}{0} = ? \text{ for } k = 0.025.$$

This shows that the gain will increase with a larger amount of positive feedback until a critical stage is reached where seemingly the gain is infinite. Of course, the output cannot be infinite, but an output voltage will develop whose frequency may bear no resemblance to the input signal. In other words, the circuit has stopped amplifying and started oscillating. The input can be removed, and the output will continue, its frequency depending upon the feedback network or the amplifier, or both.

18-2. REQUIREMENTS FOR OSCILLATION

In practice, it is not sufficient just to make $kA_v = 1$ in the equation,

$$A_{vf} = \frac{}{1 - kA_v}.$$

kA_v is made larger than 1, at which point the above equation does not hold, since the circuit is no longer an amplifier.

Since k cannot usually be greater than 1 (if k is an attenuation network), then, $kA_v \geq 1$ means that $A_v \geq 1/k$.

Say that 1.5% of the output of an amplifier is fed back positively to the input. What is the minimum gain required of the amplifier for oscillations to occur?

$$A_v \geq \frac{1}{k} = \frac{1}{0.015} = \frac{100}{1.5} = 66.6.$$

Actually, if $kA_v > 1$, the gain of the whole system will adjust itself until $kA_v = 1$.

Thus, if $A_v = 100$ and $k = 0.005$ (0.5%), $kA_v = 0.5$, and the circuit will not oscillate. But if the feedback is increased to 2%, $k = 0.02$ and the system will oscillate, since $kA_v = 2$, which meets the requirement for oscillation. For *linear oscillations*, in which the output is relatively free from distortion, kA_v should be only slightly greater than one.

Consider the diagram in Fig. 18-2, in which the amplifier causes a phase shift of 180° between V_i and V_o'. (A_v is negative as in a single stage CE amplifier.) In order to provide positive feedback, the feedback network must (a) provide a phase shift of 180° so as to provide a signal that can replace V_i, and (b) not cause more attenuation than $1/A_v$, since kA_v must be equal to or greater than 1. This condition of unity loop gain, $kA_v = 1$, is called the Barkhausen criterion for oscillation and may be met by k and A_v being both positive or both negative. When this criterion is satisfied, S-2 can be closed and V_i removed, and the output will continue. Its frequency may be different, depending upon the particular circuit.

If the requirement for oscillation, $kA_v \geq 1$, is met at only one frequency (due to frequency selective components in the feedback network or amplifier), sinusoidal oscillations will result,

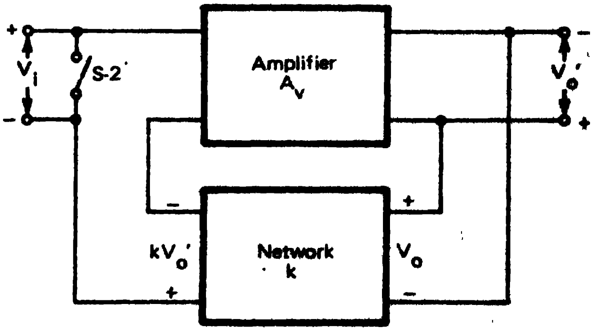


Fig. 18-2. Block diagram of an oscillator in which both A_v and k are negative.

and we have a sine wave oscillator. If the requirement is met over a band of frequencies, then the output may be nonsinusoidal in waveform, such as a square wave, etc.

18-3. THE PHASE-SHIFT OSCILLATOR

The phase-shift oscillator exemplifies all the oscillator principles discussed so far. One version, shown in Fig. 18-3, consists of an amplifier stage and a feedback network to provide an additional 180° phase shift.

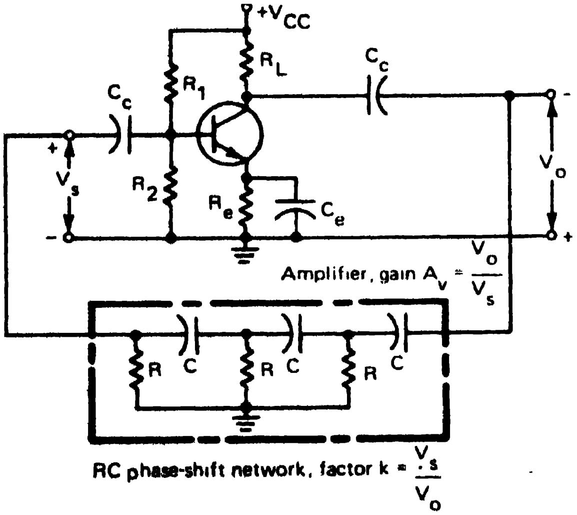

+V_{CC}
R_L
C_c
C_c
R_1
+
V_s
R_2
R_e
C_e
-
V_o
-
+
Amplifier, gain $A_v = \frac{V_o}{V_s}$
R
C
R
C
R
C
RC phase-shift network, factor $k = \frac{V_s}{V_o}$

Fig. 18-3. One version of an RC phase-shift oscillator.

The CE introduces a 180° phase shift, approximately, depending upon the frequency of operation. The RC phase-shift network must provide 180°, or an average 60° phase shift per leg. How does an RC series circuit do this? Refer to Fig. 18-4.

Consider I to flow through both R and C. Then, using I as the reference vector, V_o is in phase with I, while V_C, the voltage across the capacitor, is 90° behind I. But V_i is the sum of V_C and V_o. Hence V_o is θ degrees ahead of V_i and represents a phase shift of θ degrees. To find how much phase shift, recall:

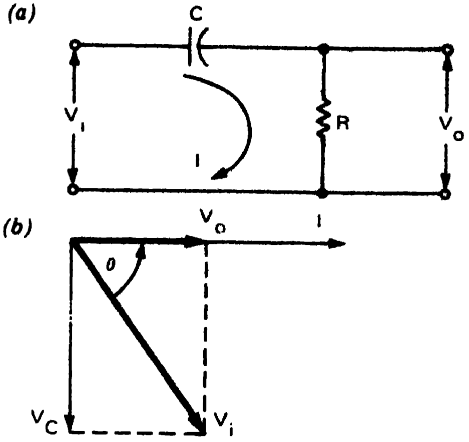


Fig. 18-4. (a) Sincle RC phase-shift network. (b) Voltage and current phase reiationships.

$V_o = IR$, $V_C = IX_C$, and $\tan\theta = \frac{V_C}{V_o}$.

$$\operatorname{Tan}\theta = \frac{IX_C}{IR} = \frac{1}{2\pi fCR}.$$

$$\therefore f = \frac{1}{2\pi RC\tan\theta}. \qquad (18\text{-}2)$$

If there are three sections, each must give, approximately, $\theta = 60°$. But $\tan 60° = \sqrt{3} \approx 1.73$.

(18-3)

This gives the approximate frequency of oscillation of a phase-shift oscillator with three sections.

For example, if R = 220 K ohms, C = 500 pF, then

$$f = \frac{}{2\pi\sqrt{3} \times 220 \times 10^3 \times 500 \times 10^{-12}}$$

$$= 835 \text{ Hz}.$$

By changing the values of R or C simultaneously in each section, the frequency of oscillation can be changed from as low as 20 Hz to as much as 300 kHz.

The above phase discussion ignored the additional current I that flows through C for the other sections, so that V_C is actually larger than the value indicated, which means that f is smaller. More accurately

$$f = \frac{}{2\pi\sqrt{6}RC}, \qquad (18\text{-}4)$$

so the above example would yield
$f = 835/\sqrt{2} = 590$ Hz.

Note that there is only one frequency at which the overall shift is 180° (or whatever is required), so that a sine-wave output is obtained.

This output will be relatively free from distortion if the gain of the amplifier slightly exceeds the loss in the RC network to make the loop gain a little over 1. Since it can be shown that a three-section network causes the output to be 1/29 of the input, the gain of the amplifier must exceed 29, and the transistor must be chosen accordingly.

18-4. ANOTHER VERSION OF THE PHASE-SHIFT OSCILLATOR

Equation 18-4 was based on no loading of the last RC leg by the input resistance to the transistor or the biasing resistors, R_1 and R_2. However, the relatively low value of $R_1 \approx h_{ie}$ will affect the frequency of operation. One way to avoid this loading is to add an emitter-follower stage to buffer the feedback network from the input transistor. It also provides a convenient means of deriving the output. See Fig. 18-5.

Note the variable control in the emitter lead of Q1. This permits the introduction of some negative feedback to control the gain of the amplifier and provide conditions for linear oscillations.

One disadvantage of the phase-shift oscillator is apparent when one tries to build a variable frequency circuit. It is difficult to adjust equally the three capacitor values simultaneously, although such variable-frequency circuits have been built. (The resistive elements of the RC leg are not adjusted. This is to avoid changing the overall impedance of the phase-shift network which would affect the loop gain, kA_v, and the amplitude of oscillation.)

18-5. WIEN BRIDGE OSCILLATOR

A circuit that lends itself more readily to variable frequency operation is the Wien bridge oscillator shown in Fig. 18-6.

The two-stage CE amplifier provides an approximate 360° phase shift, so the frequency selective network has no need to introduce any additional phase shift. The conditions for V_2 to be in phase with V_1 are met when $X_C = R$ or,

$$2\pi RC . \qquad (18\text{-}5)$$

Under these conditions, it can be shown that $V_2 = 1/3\ V_1$, so the attenuation is not as high as in the phase-shift network. (See Problem 18-7.)

Variable frequency operation requires the simultaneous adjustment of the two capacitors, an easier task than in the phase-shift oscillator. Although there are a number of bridge oscillators, the Wien bridge oscillator is capable of oscillating at frequencies from 20 Hz to 200 kHz, with a distortion of less than one percent.

18-6. RESONANT CIRCUIT OSCILLATORS

Oscillators using resonant LC tank circuits are most often used for sources of RF (radio

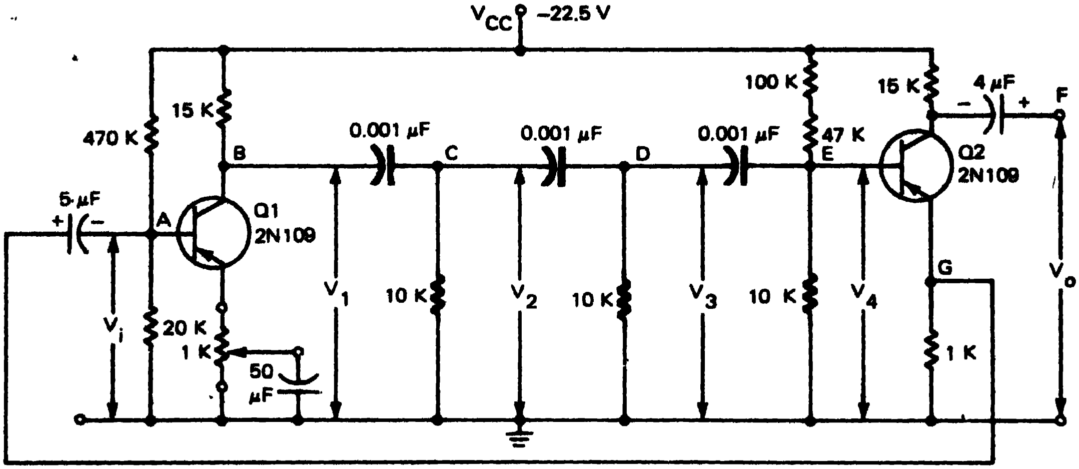

V_CC
−22.5 V
470 K
15 K
0.001 µF
0.001 µF
0.001 µF
100 K
47 K
15 K
4 µF
−
+
F
B
C
D
E
Q1
2N109
Q2
2N109
5 µF
+
−
A
V_i
20 K
1 K
50 µF
V_1
V_2
V_3
V_4
10 K
10 K
10 K
G
1 K
V_o

Fig. 18-5. Phase-shift oscillator with emitter-follower to reduce loading.

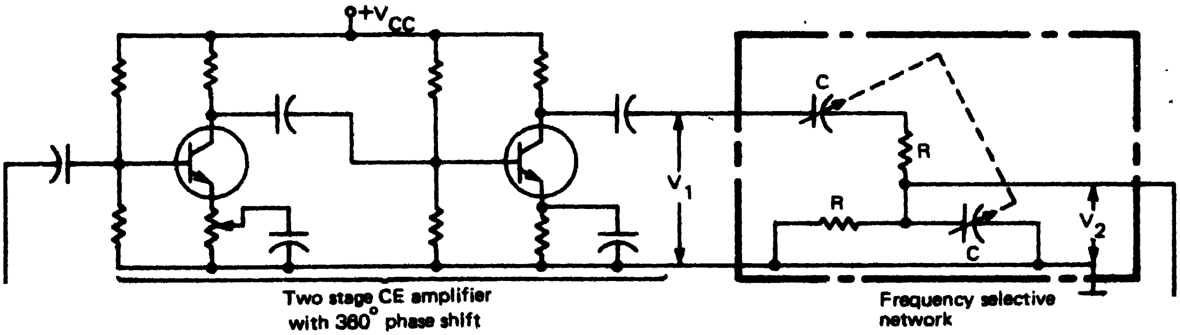

+V_CC
V_1
C
R
R
C
V_2
Two stage CE amplifier with 360° phase shift
Frequency selective network

Fig. 18-6. Wien bridge oscillator.

frequency) energy and would be covered in more detail in a communications course. However, two of the better known circuits will be covered briefly.

The Colpitts Oscillator

This can be recognized by the two series-connected capacitors, C_1 and C_2, which often are provided by a split-stator adjustable capacitor with the rotor grounded. In this way, a stable circuit can be built, which may be used as a local oscillator in a superheterodyne radio receiver. An example is shown in Fig. 18-7

This can be shown as an amplifier (CE), driving a tank circuit with a feedback path, as illustrated in Fig. 18-8.

At low frequencies, V_o is 180° out of phase with V_i, but at higher frequencies there will be a phase shift across the LC network, so that at a given frequency V_o is in phase with V_i.

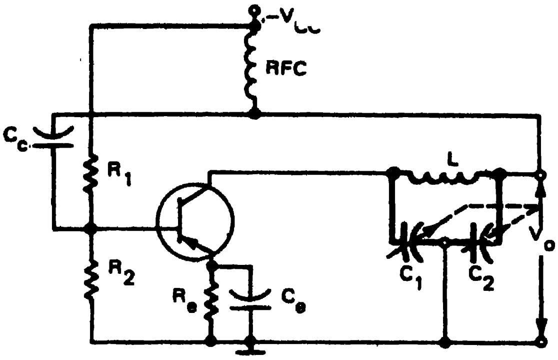


Fig. 18-7. Colpitts oscillator. Note the split capacitor.

This output is fed back through C_c (a short circuit at high frequencies) to the input V_i, whereas the RFC (radio frequency choke) has a high impedance to the RF and keeps it from going to ground through the power supply. Collector current can still be supplied through the RFC and L, and bias is maintained by R_1, R_2. Usually, collector current will flow almost continually to give near-class-A operation. However, if large power output and high efficiency are essential, class C operation will be chosen, in which collector current flows only in pulses. The output voltage, however, will be sinusoidal, because the resonant tank circuit produces a high impedance only to the fundamental component in the current.

The frequency of oscillation is given by

$$\frac{1}{2\pi\sqrt{LC}}, \quad C = \frac{C_1 C_2}{C_1 + C_2} \qquad (18\text{-}6)$$

Although the output may be taken from the oscillator, as shown by V_o in Fig. 18-8, it is usually obtained by inductive coupling to L. For some applications, a crystal may be used to replace the inductance L in the tank circuit. The resultant frequency will depend upon the natural resonant frequency of the quartz piezoelectric crystal and can be as accurately known as 1 part in 100 million, whereas the frequency of the previous circuit can drift and not be as stable over a given period of time.

The Hartley Oscillator

A Hartley oscillator can be recognized by the presence of a tapped coil and a capacitor across it, as shown in Fig. 18-9.

The voltage developed across the tank circuit is fed back to the base-emitter junction through the 0.02 μF capacitor and the variable

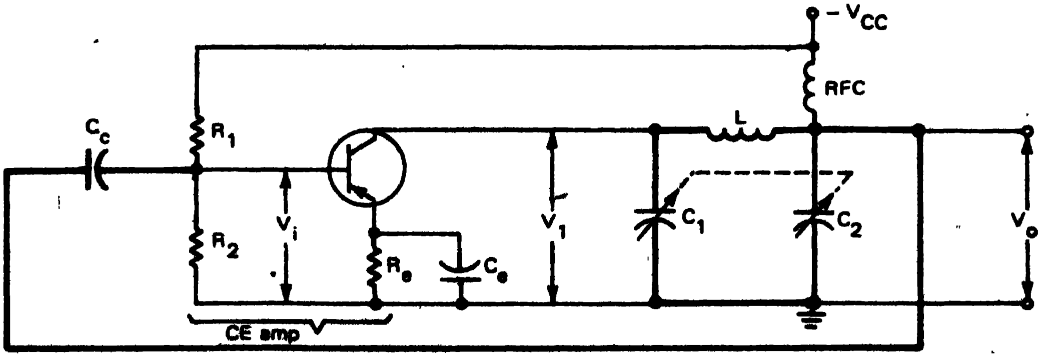


Fig. 18-8. Colpitts oscillator redrawn to show the tank circuit and feedback path.

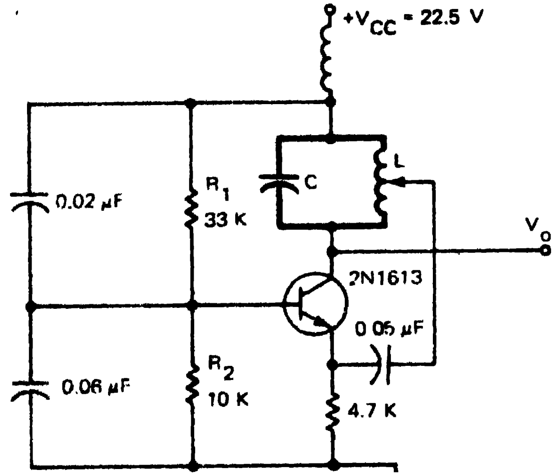


Fig. 18-9. Hartley oscillator.

tap on the coil, which produces the correct phase relationship. Resistors R_1 and R_2 bias the circuit into class C operation with pulses of current flowing through the parallel resonant circuit at a rate controlled by the resonant frequency of the tank circuit,

$$f = \frac{1}{2\pi\sqrt{LC}}, \qquad (18\text{-}7)$$

where L is the total inductance of the coil.

18-7. THE MAXIMUM FREQUENCY OF OSCILLATION OF A TRANSISTOR*

We have assumed thus far that the frequency of oscillation was determined entirely by the elements external to the transistor. At increased frequencies of oscillation, however, the capacitive effects of the transistor must be taken into account, for these will affect the necessary phase relationships. It may be shown, using the high-frequency model for the transistor, that there is an upper limit to the frequency of oscillation for a given transistor, beyond which the Barkhausen criterion cannot be satisfied. An approximation to this maximum frequency, f_{max}, is given by

$$f_{max} \approx \sqrt{\;8\pi r_{bb'} C_{b'c}} \qquad (18\text{-}8)$$

where $r_{bb'}$ is the base-spreading resistance, $C_{b'c}$ the collector-base junction capacitance, and f_T is the transistor's short-circuit, common-emitter current gain-bandwidth product.[**] Once again, we see the importance of the high-frequency characteristic, f_T. For example, for the 2N2218 silicon transistor operating at a collector current of 1.5 mA, $r_{bb'} = 140$ ohms, $C_{b'c} = 4$ pF, and $f_T = 300$ MHz. In this case, $f_{max} = 146$ MHz and is evidently lower than f_T, although this is not necessarily the case for all transistors. It depends upon the relative size of the time constant given by $r_{bb'}$ $C_{b'c}$, and f_T, so that if

$$r_{bb'} C_{b'c} \leq 8\pi f_\pi \qquad (18\text{-}9)$$

then $f_{max} \geq f_T$.

18-8. DECOUPLING FILTERS

Having examined ways of making oscillators work, let us briefly look at how undesired oscillations in an amplifier can be prevented.

In Chapters 15 and 16, we discussed cascaded CE stages, which had their collector load resistors tied to the same power supply. But we have seen how two such stages can provide 360° phase shift between input and output, so

---

[*] This section may be omitted with no loss of continuity.

[**] Gibbons, J., *Semiconductor Electronics*, McGraw-Hill, New York, 1966, Section 18-3.

is there not a possibility of positive feedback occurring, through the common power supply connection, to cause oscillation? The answer to this is in the affirmative, as can be seen by referring to Fig. 18-10.

To clarify the diagram, the load resistor for Q1 and the biasing resistors for Q2 have not been shown. The power supply has been shown to have an internal resistance, R_{INT}, which is the culprit in our problem. That is, the collector current I_{c2} flowing through the power supply will develop an ac voltage, V_T.

$$V_T = I_{c2}R_{INT}. \tag{18-10}$$

This small ac voltage at the terminals of the power supply will be able to provide feedback current, I_f, which is approximately given by

$$I_f \approx \overline{R_{L1}' + h_{ie1}} \quad \frac{I_{c2}\,R_{INT}}{R_{L1}' + h_{ie1}} \tag{18-11}$$

where $R_{L1}' = R_L \| R_1$. If R_2, the other biasing resistor sometimes used from base to ground, is involved, a slightly different expression results, but for the sake of simplicity and to understand the concept, consider it to be large compared with h_{ie1}.

Now if this feedback current, I_f, is equal to or greater than the original signal input current, I_b, (and of the correct phase, which two stages may well provide at some frequency), a condition of self-sustained oscillations exists, and an output will exist with no input. The oscillations are often at some low frequency of a few cycles per second, and in addition to the normal audio. This condition is referred to as *motor-boating* because of the "putt-putt-putt" sound produced.

Now we know the expression that relates I_{c2} to I_b. It is the overall current gain = $(-A_{i1}) \times (-A_{i2})$, taking into account any current division as covered in Section 15-3. Thus, for oscillations to occur,

$$\frac{I_{c2}}{I_{b1}} = \frac{I_{c2}}{I_f} = \frac{R_{L1}' + h_{ie1}}{R_{INT}} \quad A_{i1} \times A_{i2}. \tag{18-12}$$

Or, solving for R_{INT}, the maximum value of internal resistance in the power supply that can be tolerated before oscillations will occur, is given by

$$R_{INT} \leq \frac{R_{L1}' + h_{ie1}}{A_{i1} \times A_{i2}} \tag{18-13}$$

As an example, using typical values from previous examples, assume $A_{i1} = A_{i2} = 50$, $R_L = 2$ K ohms, $R_1 = 2$ K ohms, $h_{ie1} = 1$ Kohm.

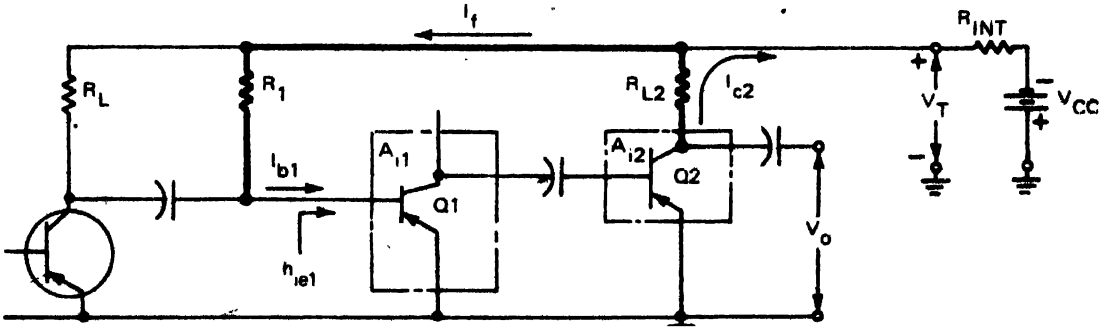


Fig. 18-10. Showing how two CE stages, driven from the same power supply, provide a positive feedback path through the power supply leads.

Then $R_{INT} = 0.8$ ohm, and so, for safety, the maximum value of R_{INT} should not exceed 0.5 ohm. When it is realized that the power supply for the amplifier usually consists of some rectifier and filter circuit of the type studied in Chapters 2 and 3, it is not hard to see how this is difficult to come by. Only power supplies having some type of regulating circuit (see Chapter 23) will be able to approach this.

The problem is usually overcome by the use of decoupling filters, which reduce the feedback by some desired amount. This is shown in Fig. 18-11 for the circuit of Fig. 18-10.

The combination of R_F and C_F form a filter, in which, if the reactance of C_F is small compared with R_F, even at the lowest frequency of operation, the ac voltage V_F will be considerably smaller than V_T. For example, if the feedback is to be reduced to approximately 10% of its original value, $X_{CF} = 0.1\ R_F$. However, the dc current from the power supply to the earlier stages must pass through each series decoupling resistor, R_F, so that the value of R_F will depend upon the value of voltage drop that can be tolerated.

Example 18-1. If the last stage operates at $V_{CC} = 22.5$ V, and the earlier stage before the decoupling filter requires 20 V when the total current through R_F is 2 mA, determine the value of C_F to reduce the feedback to 10%. The lowest frequency of operation is to be 50 Hz.

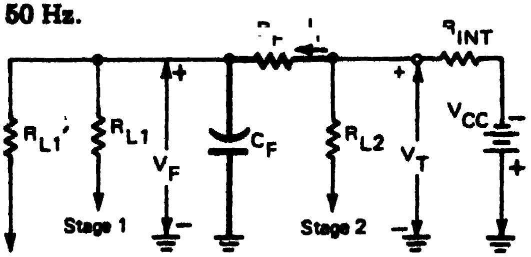


Fig. 18-11. Decoupling filter consisting of R_F and C_F to reduce the positive feedback current in the power supply leads.

Solution

$$R_F = \frac{V_{CC} - V}{I} = \frac{22.5 - 20}{2 \times 10^{-3}} = 1.25 \text{ K ohms.}$$

$$X_{CF} = 0.1\ R_F = 125 \text{ ohms}$$

$$C_F = \frac{1}{2\pi \times 50 \times 125} = 25$$

The effect of the decoupling filters is to isolate the operation of one stage from another as far as having a common supply is concerned.

18-9. THE TRANSISTOR AS A SWITCH

Since a multivibrator involves two transistors, which alternate between conduction and cutoff, we should look at the switching characteristics of a transistor and the time required to change from one state to the other.

Consider the circuit in Fig. 18-12(a), where a step input, v_i, is applied at the base of the transistor. The load line for V_{CC} and R_L has been drawn on the transistor's characteristic curves in Fig. 18-12(b), where it is seen that two extreme states exist.

The Open Switch, R_{off}

With $v_i = 0$ V, $I_B = 0$, the transistor is cutoff, no current flows through R_L except the leakage current I_{CEO}, and the "switch" is in the off position. Note that in this state the voltage from collector-to-emitter is maximum and equal to V_{CC}, since there is negligible voltage drop across R_L when $I_C \approx 0$. An example of the resistance represented by the transistor in the off state can be calculated using $I_{CEO} = 0.2$ mA, $V_{CE} = 12$ V.

$$R_{off} = \frac{V_{CE}}{I_{CEO}} = \frac{12}{0.2 \times 10^{-3}} = 60 \text{ K ohms.}$$

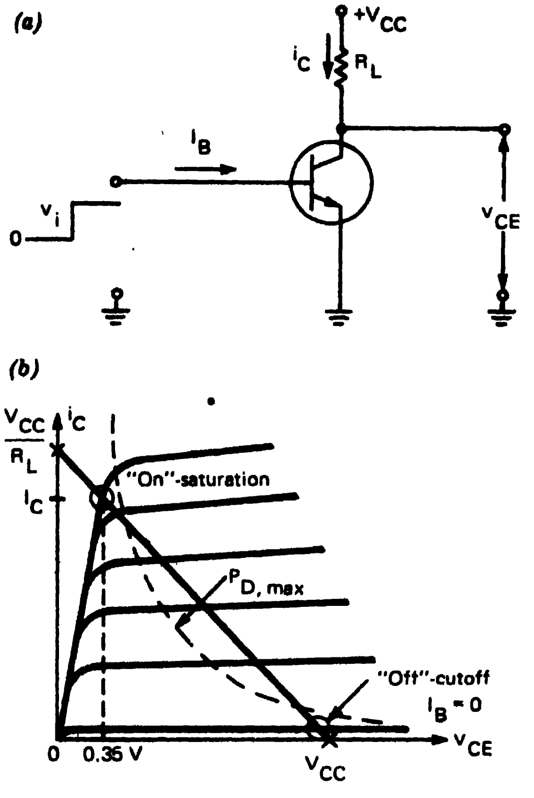


Fig. 18-12. (a) Transistor as a switch with a step input of voltage. (b) Load line on characteristic curves, showing the two extreme states.

Where the leakage current of 0.2 mA is objectionable, it may be reduced by reverse biasing the base-emitter junction, as illustrated in Fig. 18-13.

This will reduce the leakage current down to the cutoff current, I_{CO}. In a typical switching transistor, such as a 2N404, this would be approximately 5 μA.

$$R_{off} = \frac{V_{CE}}{I_{CO}} = \frac{12}{5 \times 10^{-6}} = 2.4 \text{ M ohms.}$$

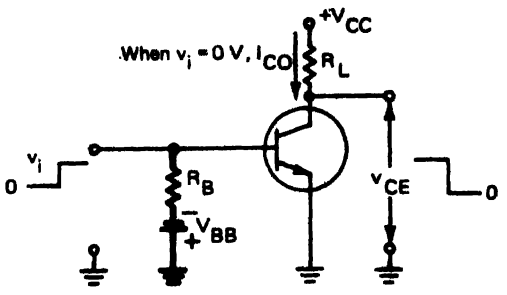


Fig. 18-13. Transistor switch biased off to reduce leakage current to I_{CO}.

In a silicon transistor, this off resistance would be even larger under reverse-bias conditions. Say $I_{CO} = 25$ nA.

$$R_{off} = \frac{V_{CE}}{I_{CO}} = \frac{12}{25 \times 10^{-9}} = 500 \text{ M ohms.}$$

Certainly one could think of the transistor switch as being open with this much resistance. One disadvantage of using reverse bias is the longer time it requires to switch the transistor on when v_i is applied. This will be referred to again later.

The Closed Switch, R_{on}

Now consider applying v_i to turn the "switch" on by applying base current I_B. The amount of base current necessary to do this can be obtained from the load line. Usually, more current than the minimum base current to just cause saturation is used to switch the transistor. This will cause the transistor to switch on quickly when driven hard into saturation. However, if the transistor is in the heavily saturated state, it will take longer for it to switch off again when v_i is reduced to zero.

The resistance of the transistor in the on state depends upon the collector current I_C, but a typical value of saturated collector-to-emitter voltage is 0.35 V at $I_C = 200$ mA, or 0.2 V at lower I_C.

$$R = \frac{V_{CE}}{I_C} = \frac{0.35}{200 \times 10^{-3}} = 1.75 \text{ ohms.}$$

Note that V_{CE} is almost zero, and the ground at the emitter has now effectively been transferred to the collector, so that the full V_{CC} voltage has been placed across R_L.

In this sense, the transistor switch acts as an inverter, since when v_i is zero, v_{CE} is maximum at V_{CC}; and when v_i is maximum, v_{CE} is minimum near 0 V. That is, the output is the inverse of the input, although the voltage levels will not usually be the same.

If the transistor is reverse-biased, it will take a larger signal v_i to switch the transistor on than with no reverse bias, since the on signal must "swamp out" the reverse voltage.

It should be noted that the load line in Fig. 18-12(b) passes through a region in which the maximum dissipation of the transistor is being exceeded. However, since the only two *quiescent* states for the switch are in regions of permissible dissipation, this operation is acceptable if the switching from one state to the other is rapid. This would not be a good idea if the transistor were operating at a speed of 60 Hz for example. Then the load line should be completely within the maximum dissipation hyperbola. The advantage of the operation shown in Fig. 18-12(b) is that higher values of collector current can be used than if the load line remained within the maximum dissipation curve at all times. However, the disadvantage is a higher collector-to-emitter saturation voltage.

18-10. SWITCHING TIMES IN A TRANSISTOR

Due to the capacitive effects in the transistor, the application of a square wave input to the base-emitter junction is not followed *directly* by a change in the output voltage. There is some delay. In order to measure how quickly the output changes, it is necessary to define what is meant by the rise time and fall time. The rise time, t_r, is the time required for the output to change from 10% of its final value to 90% of its final value. This removes the confusion where the "corners" are rounded and where it is difficult to determine the beginning and the end of the change in level. Consider the application of a positive square wave (1 μsec pulse) to the input of a transistor as in Fig. 18-13 and the resulting output shown in Fig. 18-14.

For a high-speed switching transistor, such as a silicon 2N708, there is a *delay* of t_d after applying v_{BE} before i_C starts to rise. The *rise* time to reach 90% of the final value is t_r. The sum of these two is the turn-on time, $T_{on} = t_d + t_r = 40$ nsec. (1 nsec = 1 nanosec = 1×10^{-9} sec.) When v_{BE} goes to zero, there is a delay before i_C starts to change due to the charge stored in the transistor being "cleaned out"; this is referred to as the charge *storage* time, $t_s = 25$ nsec. The *fall* time to reach 10% of the initial value, $t_f = 50$ nsec. The sum of these two is the turn-off time, $T_{off} = t_s + t_f = 75$ nsec. It is the rise and fall times of a transistor that limit the frequency at which the transistor can be switched. However, if the 2N708 were used in a circuit at a switching speed of just a few kHz, the above on and off times would be negligible. Then the output waveform would be determined by other elements in the circuit, exterior to the transistor, as shown in the next section.

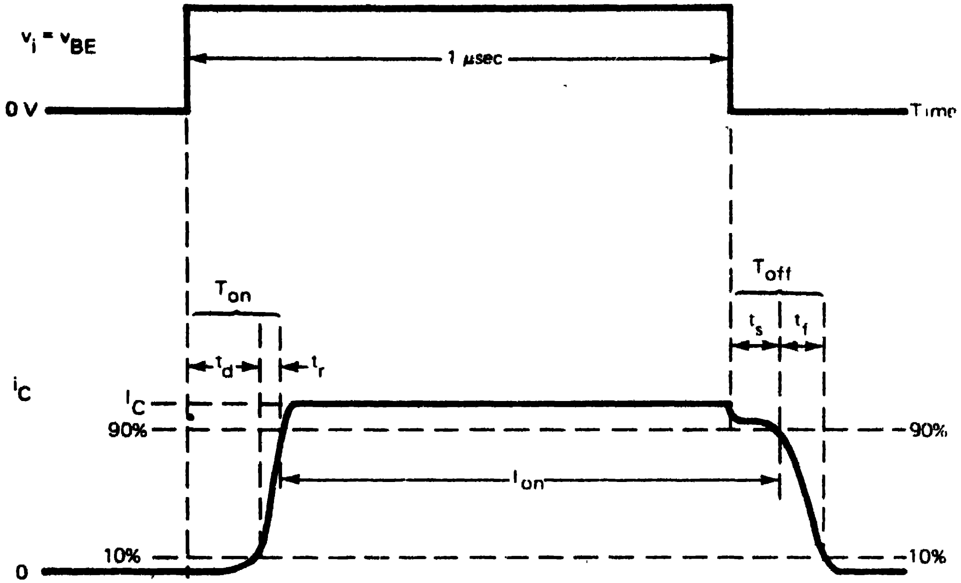


Fig. 18-14. Showing the input square wave, v_i, the resulting output waveform, i_C, and the time delays involved when the transistor is driven into heavy saturation.

In the above example, if v_{BE} is 1 μsec wide, how long is I_{on}? $I_{on} = 1000$ nsec $- T_{on} + t_s = 1000 - 40 + 25 = 985$ nsec $= 0.985$ μsec.

It should be noted that the rise and fall times depend upon the transistor's high-frequency characteristic, f_T, as discussed in Chapter 16.

18-11. MULTIVIBRATORS

There are three groups of multivibrators. The one that we shall analyze here is of the free-running, or astable, type. That is, it has two quasi-stable (half-stable) states. (Some refer to it as having no stable states which is the real meaning of *astable*.) It switches back and forth from one state to the other, remaining in each for a time dependent upon the discharging of a capacitive circuit. It is then an oscillatory circuit, since it requires no external triggering. (It does of course need a source of power.)

The other two groups of multivibrators are the monostable and bistable circuits. As the name suggests, the first (also called a one-sho has one stable state which it returns to after the application of a trigger input. It supplies a single pulse of a desired duration for every trigger input. The bistable multivibrator has two stable states in which it can come to rest after some input signal is applied. It is also known as a flip-flop, because of the two possible states it can assume, and is used as a digital memory device. It does not oscillate. We shall look at detailed circuits of these in discrete and integrated form in Chapter 26.

The Astable Multivibrator

Figure 18-15 shows an example of a basic symmetrical astable multivibrator in which

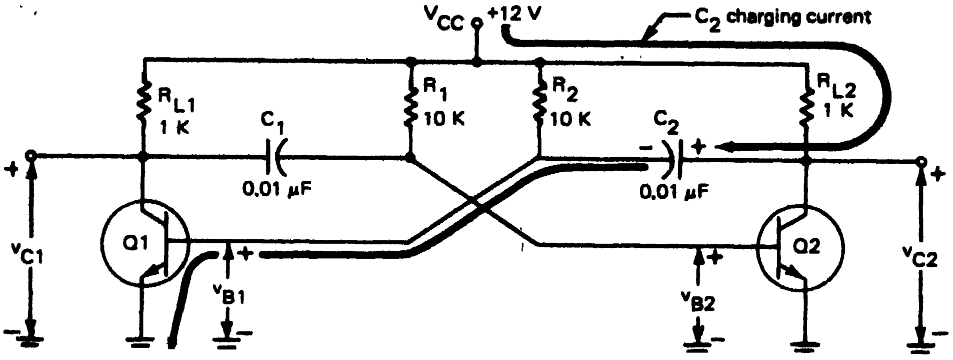


Fig. 18-15. Basic astable multivibrator.

the components in one half of the circuit are considered to be equal to their counterparts in the other half.

To analyze the circuit, we assume it is oscillating and justify that it will continue to do so:

(a) When power is applied, one transistor will conduct more than the other due to some circuit imbalance. Assume Q1 is conducting and Q2 is cutoff. Then v_{C1}, the output of Q1, is approximately 0 V, and $v_{C2} \approx 12$ V.

(b) This will permit C_2 to charge up through R_{L2} and the base-emitter junction of Q1, with the polarity shown in Fig. 18-15, to approximately 12 V.

(c) Meanwhile, C_1 had previously charged up to 12 V (when Q2 was conducting) through R_{L1}. But when Q1 conducts (the assumption made), a path is provided for C_1 to discharge through Q1 and R_1, in the manner shown in Fig. 18-16(a).

(d) The initial pulse of discharge current from C_1 through R_1 makes the base of Q2 suddenly very negative, approximately -12 V, as shown by the waveform of v_{B2} in Fig. 18-17. This causes Q2 to be cutoff, as assumed.

(e) Q1 is kept in the conducting state by base current provided from V_{CC} through R_2, so v_{B1} is slightly positive. The charging current of C_2 through R_{L2} has now ceased.

(f) The length of time for which Q2 is held off is determined by the time constant of the *discharging* of C_1 through Q1 and R_1. After approximately five time constants, Q2 is no longer cutoff but *starts* to conduct.

(g) This allows C_1 to be recharged very quickly through R_{L1} (1 K ohm) and Q2 (see Fig. 18-16(b)), causing a momentary increase in the positive voltage at the base of Q2 and aiding the turning on of Q2. The result is a fast falling edge in the waveform of v_{C2} as shown in Fig. 18-17.

(h) At the same time, C_2 is allowed to discharge through Q2 and R_2, coupling a large negative pulse to the base of Q1 (due to the voltage developed across R_2 upon the initial discharge of C_2). See Fig. 18-16(c). This turns Q1 off quickly, and the collector voltage, v_{C1}, rises *toward* +12 V. However, the voltage at the collector of Q1 is somewhat less than 12 V *at this instant* due to the voltage drop across R_{L1} occurring *while* C_1 is recharging. This is a relatively short time, but it causes a rounding of the *rising* edge in v_{C1}.

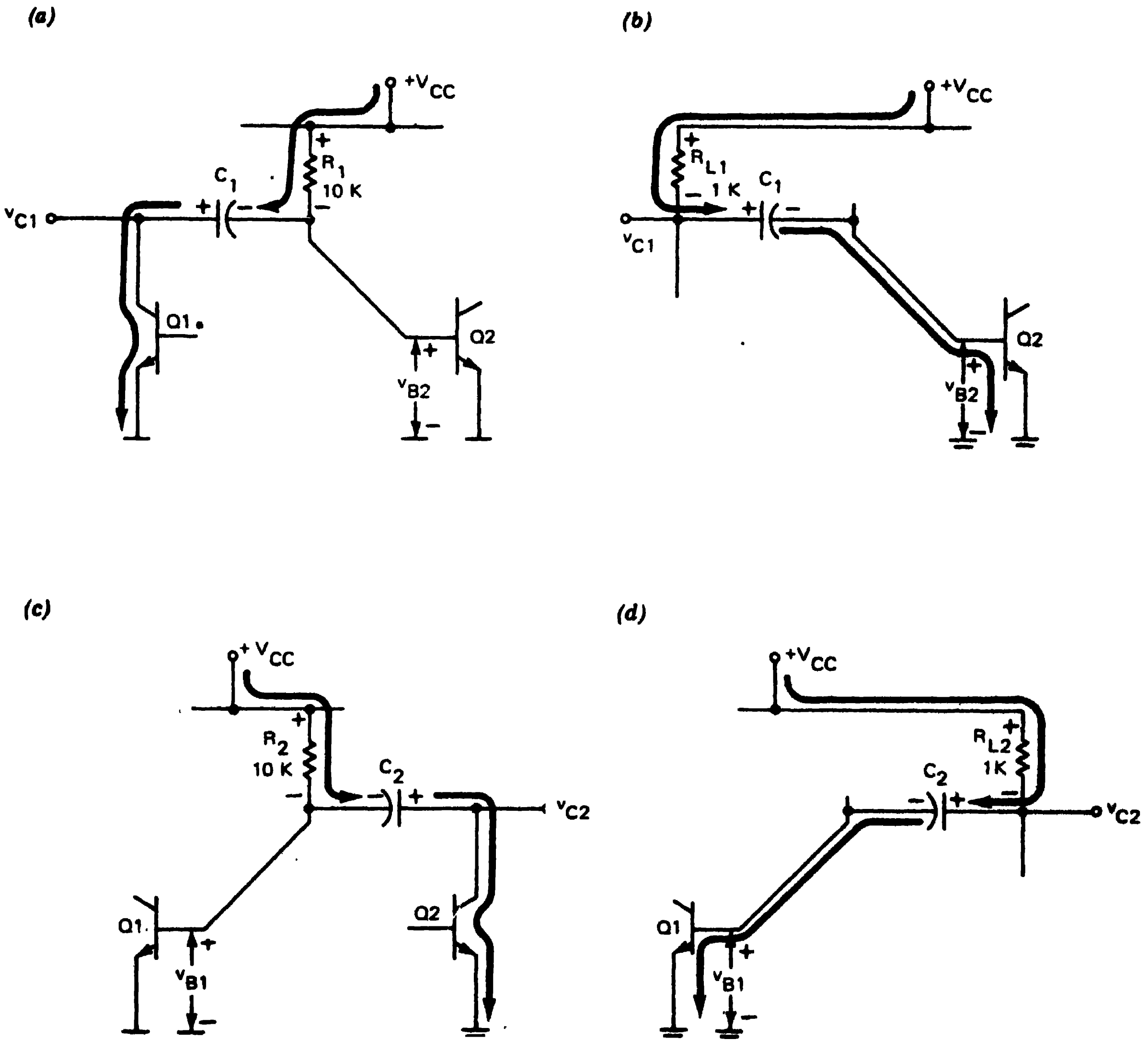

(a)
+VCC
C1
R1
10 K
VC1
Q1
Q2
VB2
(b)
+VCC
RL1
1 K
C1
VC1
Q2
VB2
(c)
+ VCC
R2
10 K
C2
VC2
Q1
Q2
VB1
(d)
+VCC
RL2
1K
C2
VC2
Q1
VB1

Fig. 18-16. Partial diagrams showing (a) The discharging path for C_1 through Q1 and R_1, and how a negative voltage is coupled to the base of Q2. (b) The recharging path for C_1 through Q2 and R_{L1}. (c) The discharging path for C_2 through Q2 and R_2. (d) The recharging path for C_2 through Q1 and R_{L2}.

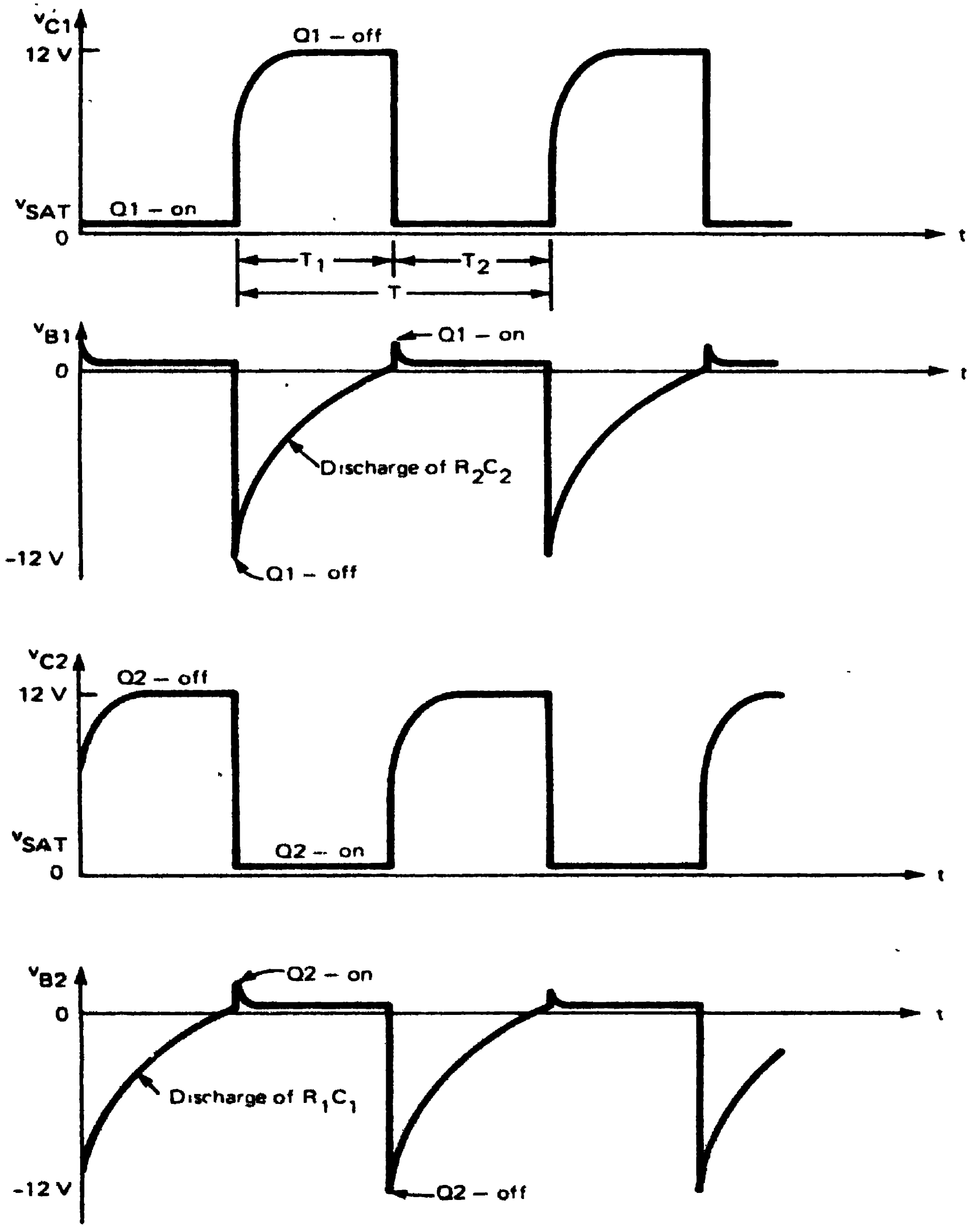

v_{C1}
12 V
Q1 – off
V_{SAT}
Q1 – on
0
t
T_1
T_2
T
v_{B1}
Q1 – on
0
t
Discharge of R_2C_2
-12 V
Q1 – off
v_{C2}
Q2 – off
12 V
V_{SAT}
Q2 – on
0
t
v_{B2}
Q2 – on
0
t
Discharge of R_1C_1
-12 V
Q2 – off

Fig. 18-17. Waveforms at the base and collector of transistors Q1 and Q2 in an astable multivibrator.

(i) Q1 is held in a nonconducting state for an interval dependent upon the R_2C_2 time constant, after which Q1 begins to conduct, C_1 discharges through Q1, and C_2 recharges through Q1, etc.

It should be noted that the turning on of one transistor, Q2, and its falling voltage at the collector permits the discharging of a capacitor which drives transistor Q1 into cutoff. The rising voltage at Q1's collector "feeds back" to the base of Q2 tending to turn it on more. The process is said to be *regenerative* or cumulative, with quite fast switching resulting.

Switching Times and Frequency of Oscillation

To obtain an expression for the frequency of oscillation, note that the period of the wave depends only upon the discharge of capacitors C_1 and C_2. Consider v_{B2} during the discharge of C_2.

$$v_{B2} \quad V_{CC} - i_{C1}R_1. \tag{18-14}$$

Since the capacitor charges up to approximately V_{CC}, the initial discharge value of

$i_{C1} \quad \dfrac{V_{CC} + V_{CC}}{R_1}$, and the current decays

exponentially with a time constant of R_1C_1.

$$\therefore v_{B2} = V_{CC} - 2V_{CC}e^{\left(\frac{-t}{R_1C_1}\right)}. \tag{18-15}$$

Let us assume that the transistor will switch when $v_{B2} = 0$ V (actually +0.5 V for a silicon transistor, or 0.2 V for germanium, which is small compared with 12 V).

$$\therefore 0 = V_{CC}\left(1 - 2e^{\left(\frac{-t}{R_1C_1}\right)}\right), \text{ or } e^{\left(\frac{t}{R_1C_1}\right)} = 2,$$

from which the off time for Q2 (and the on time for Q1) is T_2:

$$t = T_2 \quad R_1C_1 \ln 2 = 0.694R_1C_1. \tag{18-16}$$

Similarly, the off time for Q1 is T_1:

$$T_1 = 0.694R_2C_2, \tag{18-17}$$

and the total period of the wave is T:

$$T_1 + T_2 = 0.694(R_1C_1 + R_2C_2). \tag{18-18}$$

If $R_1 = R_2 = R$ and $C_1 \quad C_2 = C$, we have a symmetrical multivibrator, with the outputs at the two collectors having the same waveforms but out of phase with each other. Since $R_1C_1 \quad R_2C_2$, we have

$$T \approx 1.4RC, \tag{18-19}$$

$$\text{and } f = \frac{1}{T} \approx \frac{0.7}{RC}, \tag{18-20}$$

where T is in sec, f in Hz if R is in ohms and C in farads. For the numerical values shown in Fig. 18-15, T = 0.14 msec, f = 7 kHz. In many applications, a train of well-separated narrow pulses is required, in which case operation would be made unsymmetrical by making $R_1 \neq R_2$, by perhaps using a variable resistor. Then Eq. 18-18 would have to be used for T.

Alternatively, the frequency of oscillation may be varied by using a separate voltage source for R_1 and R_2 as shown in Fig. 18-18.

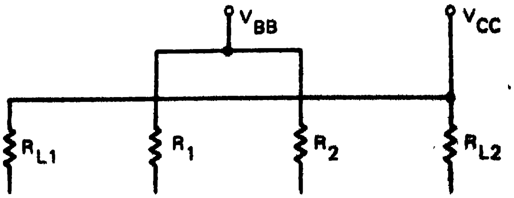


Fig. 18-18. Alternative supply voltages to vary frequency of oscillation.

The period of oscillation is then given by

$$T \approx 2RC \ln\left(1 + \frac{V_{CC}}{V}\right), \qquad (18\text{-}21)$$

where $R = R_1 = R_2$.

Minimum Value of Transistor h_{FE}

To ensure oscillation, the transistors must saturate, so that R_1 and R_2 must be chosen with the h_{FE}, the dc current gain of the transistor, in mind.

Assume Q2 is saturated, so that its collector-to-emitter voltage is approximately equal to the base-to-emitter voltage. Then the voltage drop across R_1 must equal the drop across R_{L2}.

$$I_{B2}R_1 = I_{C2}R_{L2},$$

or $R_1 = \frac{I_{C2}}{I_{B2}} R_{L2}$, but $\frac{I_{C2}}{I_{B2}} = h_{FE}$;

thus, the transistor will be saturated if

$$R_1 \leq h_{FE} R_{L2}, \qquad (18\text{-}22)$$

$$\text{or } h_{FE} \geq \frac{R}{R_L}, \qquad (18\text{-}23)$$

where $R = R_1 = R_2$,

$R_L = R_{L1} = R_{L2}$.

In the previous example, with $R = 10$ K ohms, $R_L = 1$ K ohm, a transistor with a minimum $h_{FE} = 10$ must be used.

Equation 18-20 will normally determine the value to use for R for a given frequency (for a reasonable value of C), and Eq. 18-23 will determine R_L for a given transistor.

Modifications to provide faster rise times and squarer outputs are possible, such as clamping the output voltage at some value lower than V_{CC} using Zener diodes, or diodes with auxiliary supplies. The output of a free-running multivibrator may be used to provide the switching control at the base of power transistors in a dc-to-ac inverter or operate at a higher frequency in a low-voltage dc to higher-voltage dc converter.

18-12. SUMMARY

(1) The introduction of positive feedback in an amplifier causes an increase in voltage gain. If the loop gain, $kA_v \geq 1$, the Barkhausen criterion for oscillation is satisfied, and an output remains after the input is removed.

(2) For linear oscillations, in which the output is sinusoidal with little distortion, the loop gain, kA_v, should be only slightly greater than unity.

(3) A phase-shift oscillator relies upon an RC phase-shift network to provide the necessary phase relationship between output and input to a CE amplifier. The frequency of oscillation is given by

$$f \approx \frac{1}{2\pi\sqrt{6}\,RC},$$

R in ohms, C in farads.

(4) Other versions of the phase-shift oscillator employ voltage-shunt feedback or an emitter follower to reduce loading problems.

(5) A Wien bridge oscillator uses a frequency selective bridge circuit to provide low-distortion sinusoidal oscillations at a frequency given by

$$f = 2\pi RC$$

(6) The Colpitts oscillator is an example of a resonant circuit oscillator and is characterized by a split capacitor in its tank circuit.

(7) The Hartley oscillator has a tapped coil in its tank circuit. Like the Colpitts, the Hartley

is used to generate RF (radio frequencies), and both have a frequency given by

$$f = \frac{1}{2\pi\sqrt{LC}}.$$

(8) The maximum frequency of oscillation of a transistor, f_{max}, is dependent upon the transistor's f_T as shown by the following:

$$\sqrt{\quad 8\pi r_{bb'}C_{b'c}}$$

(9) Undesired oscillations may occur in cascaded amplifier stages that operate from the same power supply by signals fed back along the power supply leads to earlier stages. The problem is usually overcome by RC decoupling filters that reduce the amount of feedback.

(10) The transistor may be used as a switch so that when it is in the saturated state its on resistance is very low, or in the cutoff state its off resistance is very high.

(11) When a transistor is switched on and off, its total turn-on time, $T_{on} = t_d + t_r$ and its turn-off time, $T_{off} = t_s + t_f$, where these times are due to capacitive delays in the transistor.

(12) An astable multivibrator is an example of a square-wave relaxation oscillator, in which the collector of each of the two transistors is capacitively coupled to the other's base.

(13) The period of oscillation of a symmetrical multivibrator depends upon the discharging time constant and is given by

$$T = \frac{1}{f} \approx 1.4RC$$

(R in ohms, C in farads).

To ensure saturation, $h_{FE} \geq \frac{R}{R_L}$, R_L = the load resistor.

QUESTIONS

Q18-1. (a) What happens as the amount of positive feedback introduced to an amplifier is increased?
(b) What is the Barkhausen criterion for oscillation?
(c) What does this imply about the relative size of k and A_v?
(d) When are linear oscillations obtained?

Q18-2. (a) What causes the output of an oscillator to be a sine wave at one frequency?
(b) How does the phase-shift oscillator get its name?
(c) By how much does the RC feedback network attenuate the input signal?
(d) What does this mean about the minimum gain the amplifier must provide?

Q18-3. How may loading of the RC phase-shift network be reduced in a phase-shift oscillator?

Q18-4. (a) What disadvantage does a phase-shift oscillator have?
(b) How is this overcome in a Wien bridge oscillator?
(c) What function does the bridge section have in the Wien bridge oscillator?
(d) Why is the bridge not used to provide phase shift?

Q18-5. How do resonant circuit or RF oscillators differ from most audio oscillators?

Q18-6. What are the distinguishing features of

(a) the Colpitts and
(b) the Hartley oscillators?

Q18-7. What factors determine the maximum frequency of oscillation of a transistor?

Q18-8. (a) Describe what problems may arise from operating two or more CE amplifiers from the same power supply.

(b) Give two ways in which the possibility of undesired oscillations may be reduced.
(c) Which method in (b) is most often used? Why?
(d) What disadvantage does the use of decoupling filters have?

Q18-9. Explain how a transistor may be thought of as a switch and as an inverter (of signals).

Q18-10. (a) What advantage is there of driving a transistor, used as a switch, into heavy saturation?
(b) What disadvantage does this have?
(c) What do the turn-on and turn-off times consist of in a switching transistor?

Q18-11. (a) What is an *astable* multivibrator?
(b) What does the half-period of oscillation of a symmetrical multivibrator depend upon?
(c) What causes the rounded leading edge in the output waveform of a multivibrator?
(d) Explain where feedback is employed in a multivibrator.
(e) Is this feedback positive or negative?

Q18-12. (a) How may the frequency of oscillation be varied in a multivibrator?
(b) How may the operation be made asymmetrical?

PROBLEMS

P18-1. An engineer designs an amplifier to have a voltage gain of 60, but when constructed it only has a gain of 50. A technician suggests that positive feedback could be used to increase the gain to the required level.
(a) How much feedback should be used to provide the desired level?
(b) What value of feedback will double the gain?
(c) What disadvantages may arise from this type of operation?
(d) What percentage of feedback will cause oscillation?

P18-2. Repeat P18-1 for a desired gain of 200 and actual gain of 150.

P18-3. A feedback network (to be used with an amplifier to provide oscillations) is tested and found to give an output of 0.0125 V with a 0.5 V input.
(a) What must be the minimum gain of the amplifier to provide oscillations?
(b) What will be the effect on the output if the amplifier's gain is twice that required by the Barkhausen criterion?
(c) What percentage *negative* feedback should be introduced to the amplifier so that *linear oscillations* result?

P18-4. A phase-shift oscillator is to use a feedback network consisting of four RC sections, similar to Fig. 18-3. Derive an approximate expression for the frequency of oscillation similar to that in Eq. 18-3.

P18-5. A three-section RC phase-shift oscillator has R = 10 K ohms and C = 0.001 μF.
(a) What is the frequency of oscillation using Eq. 18-4?
(b) If the oscillator is to be made variable, using the same values of R, what should be the tuning range of the capacitors to obtain a frequency range of 1 to 100 kHz?

P18-6. If the practical lower limit on C is 47 pF in the Wien bridge oscillator of Fig. 18-6,
(a) What should be the fixed value of R to produce an upper frequency limit of 200 kHz?
(b) What should be the tuning range of C to provide a lower frequency limit of 20 Hz?

P18-7. Use vector (phasor) algebra to obtain the expression for V_2/V_1 in Fig. 18-6.

$$\frac{V_2}{V_1} \quad \frac{RX_C}{3RX_C + j(R^2 - X_C^2)}.$$

Hence show that, if $R = X_C, V_2 = \frac{1}{3}V_1$, and the frequency of oscillation is given by

$$f = \frac{}{2\pi RC}$$

P18-8. A Colpitts oscillator is used as the local oscillator in an AM radio receiver to produce frequencies from 1 to 2 MHz.

(a) What must be the inductance of the coil if the minimum capacitance obtainable is 43 pF (C_1 and C_2)?

(b) What must be the maximum value of C to produce the necessary frequencies?

(c) What frequency would be produced if this coil were used in a Hartley oscillator circuit with C = 100 pF?

P18-9. A 40397 transistor has the following parameters specified: f_T = 50 to 80 MHz, $r_{bb'}$ = 20 to 40 ohms, C_{ob} ($C_{b'c}$) = 12 to 20 pF. Determine the *range* of the maximum frequency of oscillation this transistor may have when used in an oscillating circuit, and specify which parameter values provide the extremes.

P18-10. Two CE stages are operating from the same 30 V power supply, which has an internal resistance of 2 ohms. The first stage operates at a quiescent collector current of 0.5 mA with a bias resistor R_1 = 6 K ohms and h_{ie1} = 2 K ohms. The current gain of the first stage is 50 and of the second stage 100.

(a) Determine if the above combination is in danger of breaking into oscillations.

(b) What values should be used in a decoupling filter if the first stage can work adequately at 25 V and the feedback is to be reduced to 10% of its original value (f = 50 Hz)?

P18-11. A TIS 99 silicon transistor is used as a switch in a circuit like that in Fig. 18-12. At a base current of 3 mA, the collector current is 130 mA and V_{CE} = 0.4 V. The transistor has an I_{CBO} = 10 nA at 25°C and I_{CEO} 1 μA. If the collector supply voltage is 12 V, determine:

(a) The off resistance when $I_B = 0$.

(b) The off resistance when the base-emitter junction is reverse biased.

(c) The on resistance when I_B = 3 mA.

(d) The on resistance if I_B is increased to 4 mA, at which the collector current is 140 mA and V_{CE} = 0.3 V.

P18-12. Consider the multivibrator in Fig. 18-15.

(a) Determine the frequency of oscillation if $R_1 = R_2$ = 15 K ohms and $C_1 = C_2$ = 0.005 μF.

(b) Determine the period and frequency of oscillation if R_1 = 1 K ohm, R_2 = 20 K ohms, and C_1 = 0.01 μF, C_2 = 0.05 μF. Draw the waveform at one of the outputs.

(c) Determine the value of capacitors to use to provide a symmetrical oscillation of 50 kHz if $R_1 = R_2$ = 4.7 K ohms.

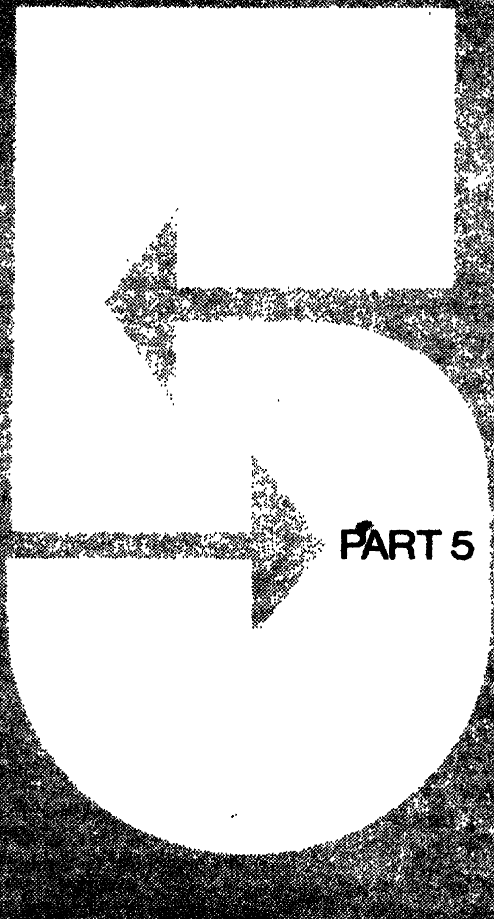
(d) Determine the value of capacitors to use to provide a train of pulses, 10 μsec wide, at a repetition rate of 10 kHz, if $R_1 = R_2$ = 1 K ohm.

(e) What must be the minimum value of load resistance to use in (a) to provide saturated operation if each transistor's h_{FE} = 40?

(f) Determine the frequency of operation for the components in (a) if the supply voltages are modified to the arrangement in Fig. 18-18 with V_{CC} = 30 V and V_{BB} = 15 V.

P18-13. (a) Verify that the expression in Eq. 18-21 reduces to that in Eq. 18-19 when $V_{BB} = V_{CC}$.

(b) What ratio of V_{CC} to V_{BB} will provide a frequency of oscillation twice that occurring when V_{BB} V_{CC}?



PART 5

TRANSISTOR POWER AMPLIFIERS

19

TRANSISTOR POWER AMPLIFIERS I: CLASS A

19-0. INTRODUCTION

The transistor audio amplifiers examined thus far have been used for the purpose of voltage amplification, the output power seldom exceeding a few milliwatts. These amplifiers were generally used with small input signal voltages and allowed the use of small-signal equivalent models for analysis purposes, whereas power amplifiers are large signal amplifiers and require a different approach.

This chapter covers the design of a class A transformer-coupled power amplifier, its efficiency and harmonic distortion. The push-pull concept is also introduced.

19-1. CLASS A DIRECT-COUPLED RESISTIVE LOAD

This is the only type of amplifier considered so far, and was shown to have a collector circuit efficiency, better called the conversion efficiency, η, of only 25%. This has been defined as follows:

$$= \frac{\text{ac power delivered to load}}{\text{dc power supplied by battery}} \times 100\%$$

$$= \frac{P_L}{P_B} \times 100\%.$$

This figure is the optimum value possible but is nearly attained in transistor amplifiers. However, where a large amount of ac power is required this means a great deal of wasted power. This power partially goes into heating up the transistor, so the second item to examine is the relation of the useful ac power to the power dissipated in the transistor, P_D.

In a class A power amplifier, maximum dissipation occurs in the transistor under quiescent conditions (no applied signal). Referring to this as $P_{D,max}$, it can be shown that

$$P_{D,max} = 2P_L. \tag{19-1}$$

where P_L is the ac power delivered to the load by an input signal that drives the transistor over a full excursion on the load line between saturation and cutoff.

If the maximum average dissipation capability of a transistor is designated as P_T, then the maximum ac power this transistor can deliver to a load is $\frac{P_T}{2}$.

Thus, if a transistor has a P_T of 150 mW for example, then maximum $P_L = \frac{P_T}{2} = \frac{150 \text{ mW}}{2}$ $= 75$ mW, and $P_B = \frac{P_L}{\eta} = \frac{75 \text{ mW}}{0.25} = 300$ mW.

Thus, a class A amplifier with a direct-coupled load can supply only one-half the amount of ac power to the load that its transistor can dissipate, and the dc power input is four times the amount of ac signal power in the load.

This can be represented by the diagrams in Fig. 19-1, in which the large "square" represents the total dc power supplied by the battery, P_B.

These diagrams clearly show that a device operated in the class A mode is cooler when delivering power to a load than when no input signal is applied. That is, the heating of the active device is at a maximum when the input signal is zero, and the heating is reduced by the amount of dc input power converted into useful ac power in the load.

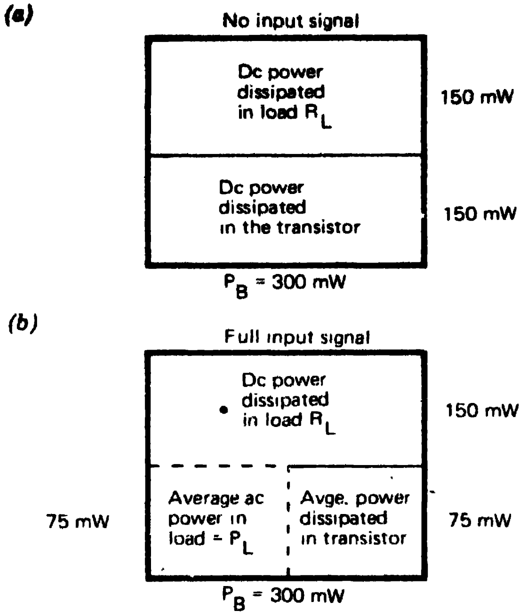

(a)
No input signal
Dc power dissipated in load R_L
150 mW
Dc power dissipated in the transistor
150 mW
P_B = 300 mW
(b)
Full input signal
Dc power dissipated in load R_L
150 mW
75 mW
Average ac power in load = P_L
Avge. power dissipated in transistor
75 mW
P_B = 300 mW

Fig. 19-1. Diagrams representing the power dissipated in the load and transistor of a directly coupled amplifier under conditions of: (a) No input signal. (b) Full input signal.

19-2: TRANSFORMER-COUPLED RESISTIVE LOAD

The conversion efficiency can be increased by a factor of two by eliminating the dc power dissipated in the load. This is accomplished by coupling the load resistor to the transistor collector circuit by means of a transformer. Since the primary of this transformer has negligible dc resistance, there is negligible dc power lost at this point. The ac power is, however, coupled magnetically across the transformer into the load R_L. This coupling method also prevents a large dc current from flowing in the load, which otherwise could be harmful if the load were a loudspeaker, since this would cause saturation of the magnetic circuit and impair the reproduction of the audio signal.

Also, the transformer is useful as an impedance-matching device to match a load resistance of a few ohms (4, 8, or 16 ohms) to the output impedance of the transistor collector circuit, which may be a few hundred ohms.

Let us consider the efficiency of a transformer-coupled-load power amplifier. Refer to Fig. 19-2.

Let I_c be the rms value of the ac current in the collector and transformer primary and V_c be the rms value of the ac voltage across the transformer primary. (Note: As far as ac is concerned, the transformer primary is in parallel with the voltage from collector to emitter, since V_{CC} is a short to ac and C_e bypasses R_e.)

Then the ac power delivered to the primary of the transformer (into the reflected resistance $R_T = \alpha^2 R_L$ where $\alpha = \frac{N_1}{N_2}$) is given by

$$P_L' = V_c I_c, \tag{19-2}$$

and the dc power delivered from the battery,

$$P_B = V_{CC} I_C. \tag{19-3}$$

From the diagram, $V_{CC} = \sqrt{2}V_c$, $V_{min} \simeq 0$

and, $I_C = \sqrt{2}I_c$, $I_{min} \simeq 0$.

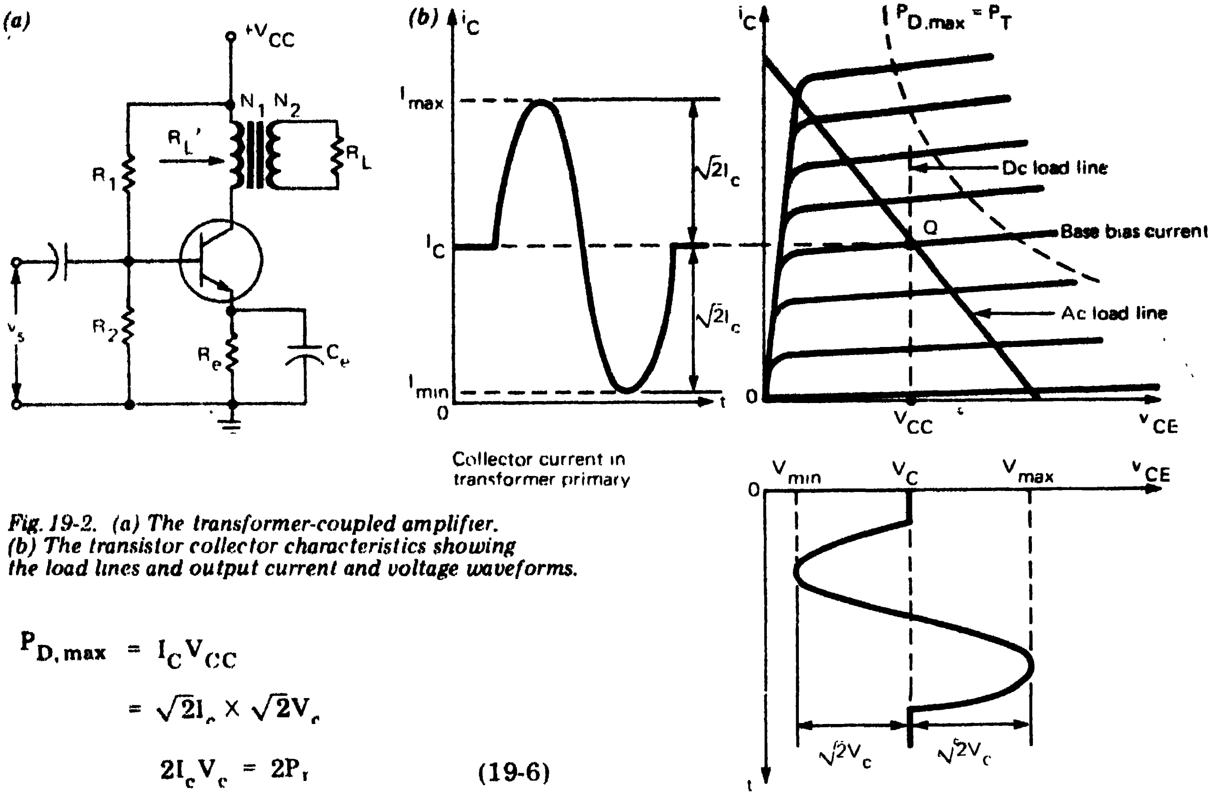
$$P_B = V_{CC}I_C = \sqrt{2}V_c \times \sqrt{2}I_c \quad 2V_cI_c. \tag{19-4}$$

$$\therefore \text{Conversion efficiency } \eta = \frac{P_L}{P_B}$$

$$\frac{V_cI_c}{2V_cI_c} \times 100\%$$

$$= 50\%. \tag{19-5}$$

This is the optimum efficiency possible for a class A transformer-coupled power amplifier. The dc input power is still constant, and the maximum dissipation in the transistor, $P_{D,max}$, occurs when the input signal is zero:



Collector current in transformer primary

Fig. 19-2. (a) The transformer-coupled amplifier. (b) The transistor collector characteristics showing the load lines and output current and voltage waveforms.

$$P_{D,max} = I_C V_{CC}$$

$$= \sqrt{2}I_c \times \sqrt{2}V_c$$

$$2I_c V_c = 2P_r \qquad (19\text{-}6)$$

Thus, if a transistor has a P_T of 150 mW, the maximum power available from the transistor, P_L', is 75 mW, and the dc power needed from the power supply battery, P_B, is $\frac{P_L'}{0.5} = \frac{75 \text{ mW}}{0.5} = 150$ mW.

Thus, for a given transistor dissipation capability, the output ac power obtainable in a load is the same no matter whether the load is direct-coupled or transformer-coupled. However, the drain on the power supply in a transformer-coupled amplifier is ideally one-half of what it is in the direct-coupled amplifier.

19-3. DESIGN THEORY

Referring to the load line and characteristic curves in Fig. 19-2(b), used earlier for the transformer-coupled amplifier, note that the dc load line is drawn vertically upward at a point where $v_{CE} = V_{CC}$. If the resistance of the transformer primary and bias stabilizing resistor R_e are included, this dc load line will slope slightly to the left. In any event, the ac load line is drawn with a slope proportional to $\frac{1}{R_L'}$ where R_L' is the resistance reflected from the secondary, R_L, into the primary. If maximum output power is the objective, this load line will be drawn tangent to the maximum dissipation hyperbola, where the dc load line intersects the hyperbola, the point of tangency determining the bias conditions. The maximum dissipation hyperbola selected will depend upon the maximum operating temperature of the transistor. Note that $V_{max} = 2V_{CC}$, which means that the load line must be drawn so as not to violate the maximum collector-emitter voltage of the transistor. This is usually done by making $V_{CC} \leq \frac{V_{CE,max}}{2}$.

Actually, if both the battery operating voltage, V_{CC}, and the desired output power, P_L', are given, then the ac load line is completely determined, as well as the operating point Q.

Using the fact that $P_{D\,max}$ is twice the ac output power, P_L', the Q point is located at

$$I_C = \frac{2P_L'}{V_{CC}} \text{ and } V_C = V_{CC}, \qquad (19\text{-}7)$$

from which, assuming Q is at the mid-point of the load line, it is left as a student exercise to show that

$$R_T \quad \frac{V_{CC}{}^2}{2P_L{}'}. \qquad (19\text{-}8)$$

This is the reflected load resistance as seen in the primary. A transformer will have to be chosen to make the actual load R_L on the secondary look like R_L' on the primary as given by this equation.

The following example will illustrate the use of the above equations in designing a power amplifier for a required output power and designated supply voltage.

19-4. POWER AMPLIFIER DESIGN

Example 19-1. Let it be required to find the value of the components in the transformer-coupled amplifier shown in Fig. 19-3. A power of 50 mW is to be delivered to the load R_L of 4 ohms, using a 2N109 germanium transistor and a dc supply of 9 volts.

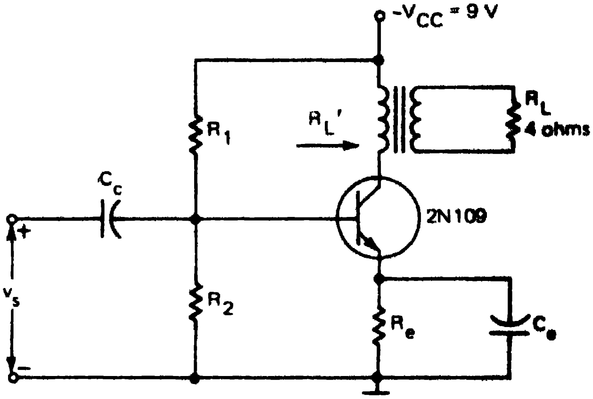


Fig. 19-3. Schematic of a transformer-coupled amplifier in which the components are to be determined for a required output power.

Solution. From the transistor manual a 2N109 has a maximum average dissipation, P_T, of 165 mW at ambient temperatures up to 25°C. We must be sure that this is not exceeded.

First, allow for the inefficiency in the output transformer. For a small transformer like this it is unlikely the efficiency will exceed 70%. Thus if P_L = 50 mW in R_L then P_L' delivered to the transformer primary (into R_L') must be

$$P_L' = \frac{50\text{ mW}}{0.70} \simeq 70\text{ mW}.$$

Checking the power dissipated in the transistor under conditions of no applied signal (this is the worst case),

$$P_{D,max} \quad 2P_L' = 2 \times 70\text{ mW} = 140\text{ mW}.$$

This is just within the 165 mW limit and indicates that this transistor is delivering practically its maximum power capability in this design.

Using $R_L' = \frac{'CC}{2P_r}$

with $V_{CC} = 9$ V, $R_L' = \quad 140 \times 10^{-3}$

$$= \frac{81 \times 10^3}{140} \simeq 580 \text{ ohms.}$$

The nearest standard transformer would be one that could match a primary impedance of 600 ohms to a load of 4 ohms. From a catalog, a TA-23 by Stancor would meet this, with a power rating of 0.15 watt (150 mW), which is easily capable of handling the 70 mW input to the transformer in this case.

Next, the bias resistors R_1, R_2, and R_e must be chosen to give the required dc operating point. This point is determined by

$V_C = V_{CC} = 9$ V,

and $I_C = \frac{2P_L'}{V_{CC}} - \frac{140 \times 10^{-3}}{} = 15.5$ mA.

The bias voltage from base to emitter determined by R_1 and R_2 must be such as to provide a collector current of 15.5 mA. Since the 2N109 is a germanium transistor and it is working near its rated dissipation, it is necessary to provide bias stabilization. This is especially important since the low dc resistance of the output transformer makes this amplifier very susceptible to thermal runaway. Stabilization is provided by R_e; let R_e be 47 ohms and check to see what this gives us.

With $I_C = 15.5$ mA flowing through R_e under quiescent conditions, the result will be a dc voltage across R_e of

$I_C R_e \quad 15.5 \times 10^{-3} \times 47 \quad 0.73\,\text{V}\ (I_C \simeq I_E)$.

This serves to reverse bias the junction from base to emitter, so that the voltage developed across R_2 must overcome this and provide about 0.27 volt from base to emitter in the forward direction, as indicated in Fig. 19-4.

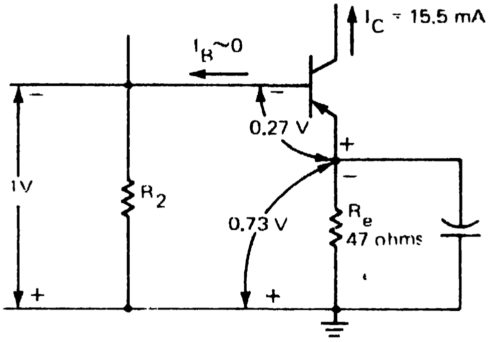


Fig. 19-4. Transistor input circuit showing bias conditions.

Thus we must develop 1 V across R_2 from the 9 V battery using R_1 to drop 8 V. In other words, using the voltage divider equation,

$$1\text{ V} = 9\text{ V} \times \frac{R_2}{R_1 + R_2}, \text{ or } \frac{1}{9} \quad \frac{R_2}{R_1 + R_2}.$$

This does not give sufficient information to solve for R_1 and R_2. However, recall that, for a low stability figure S (good stabilization), it is necessary that

$$\frac{R_1 R_2}{R_1 + R_2} \simeq 10R_e = 10 \times 47 = 470 \simeq 0.5 \text{ K ohm.}$$

Using these two equations and solving for R_1 and R_2 yields the nearest standard values of

$R_1 = 4.7$ K ohms, and $R_2 = 560$ ohms.

To complete the design, select a 100 μF capacitor @ 10 V for C_e and 10 μF for C_C to give a good low-frequency response.

By using the input characteristics for the 2N109, it may be shown that the amplifier has a power gain of 36 dB or an overall gain of 32 dB if the ac power lost in the biasing resistor, R_2, is taken into account.

19-5. HARMONIC DISTORTION

To determine the harmonic distortion for this amplifier the five-point method covered in Section 7-13 must be used. This is because the transistor is driven over wide extremes of the load line, which causes harmonics higher than the second to be significant.

Using the ac load line and the output characteristic curves of the 2N109 (not shown), the collector currents corresponding to the necessary values of base current are as follows:

$$\begin{array}{llll} I_{B,\max} = 0.34\ \text{mA} & & I_{C,\max} = 28\ \text{mA} \\ I_{B,\frac{1}{2}} = 0.255\ \text{mA} & & I_{C,\frac{1}{2}} = 24\ \text{mA} \\ I_{B,\text{bias}} = 0.17\ \text{mA} & & I_{C,Q} = 15.5\ \text{mA} \\ I_{B,-\frac{1}{2}} = 0.085\ \text{mA} & & I_{C,-\frac{1}{2}} = 10\ \text{mA} \\ I_{B,\min} = 0 & & I_{C,\min} = 0. \end{array} \tag{19-9}$$

Using these values of I_C, the following components in the output are determined:

$$A_o = \frac{1}{6}(I_{\max} + 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} + I_{\min}) - I_Q$$
$$= 0.5\ \text{mA}.$$

This is the additional dc current that flows when the signal is applied, causing the average collector current to increase to 16 mA.

$$A_1 = \frac{1}{3}(I_{\max} + I_{\frac{1}{2}} - I_{-\frac{1}{2}} - I_{\min}) = 14\ \text{mA}.$$

This is the amplitude of the fundamental, and in this case is the same as $\dfrac{I_{C,\max} - I_{C,\min}}{2}$.

$$A_2 = \frac{1}{4}(I_{\max} - 2I_Q + I_{\min})$$
$$= -0.75\ \text{mA}.$$

This is the amplitude of the second harmonic.

$$A_3 = \frac{1}{6}(I_{\max} - 2I_{\frac{1}{2}} + 2I_{-\frac{1}{2}} - I_{\min}) = 0.$$

In this case there is apparently no third harmonic.

$$A_4 = \frac{1}{12}(I_{\max} - 4I_{\frac{1}{2}} + 6I_Q - 4I_{-\frac{1}{2}} + I_{\min})$$
$$= -1.25\ \text{mA}.$$

The fourth harmonic is uncommonly large.

$$\text{Thus, } D_2 = \frac{|A_2|}{|A_1|} = \frac{0.75}{14} \times 100\% = 5.3\%,$$

$$D_3 = \frac{|A_3|}{|A_1|} = \frac{0}{14} \times 100\% = 0\%,$$

$$D_4 = \frac{|A_4|}{|A_1|} = \frac{1.25}{14} \times 100\% = 9\%,$$

and the total harmonic distortion is

$$D = \sqrt{5.3^2 + 0 + 9^2} = 10.4\%.$$

This is typical of the distortion in this type of amplifier, although it is more usual for the distortion to decrease with increase in harmonic frequency. This calculation of distortion has been made on the assumption that the source resistance is large, causing the base excitation current to be sinusoidal. A reduced amount of distortion will result if the source resistance is low compared with the input resistance to the transistor, as shown in Section 11-6. This will allow the base-to-emitter voltage to be sinusoidal and will require that the collector current be determined (for harmonic distortion calculation) from a consideration of base-emitter voltage drive rather than base current drive, as was done here. This will of course necessitate the use of the input characteristic curve of the transistor.

From a listening standpoint, the distortion at higher harmonics is more objectionable than at lower harmonics, so that the presence of 9% fourth harmonic distortion is far more objectionable than 5% (or even 9%) of second harmonic distortion.

19-6. POWER OUTPUT

It was stated that the effective load resistance seen by the collector in Fig. 19-3 was $R_L' =$ 580 ohms. Applying the equation for output power at the fundamental frequency,

$$P_1 = \frac{A_1{}^2 R_L'}{2} = \frac{(14 \times 10^{-3})^2 \times 580}{2} = 57 \text{ mW}.$$

This compares with

$$\frac{(I_{C,\max} - I_{C,\min})}{2\sqrt{2}} \times \frac{(V_{CE,\max} - V_{CE,\min})}{2\sqrt{2}}$$

$$\frac{28 - 0}{2\sqrt{2}}\ \frac{15.5 - 1.9}{2\sqrt{2}} = 47 \text{ mW}$$

using the voltage values from the graph (not shown), which shows the importance of using the correct equation. Actually we wanted 70 mW into the primary of the transformer (R_L') to obtain 50 mW in R_L on the secondary. The above calculation of 57 mW suggests we will be low in output power, but it can probably be increased by adjusting the bias point and input drive.

19-7. VARIATION OF OUTPUT POWER WITH LOAD

The way in which the power delivered to the load (and also the distortion) varies with the value of the load seen at the transformer primary is shown in Fig. 19-5 for a typical transistor amplifier.

It is seen from the curves that the load that provides minimum total harmonic distortion is different from the load that yields maximum power output. Fortunately, the power curve is fairly flat in the region of maximum power, so that a well-designed amplifier would work into a load that gives slightly less than maximum output power, but a significantly lower total distortion.

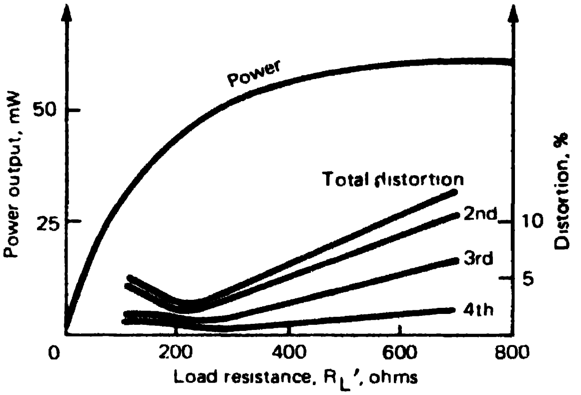


Fig. 19-5. Graph showing how output power and distortion vary with load resistance in a transformer-coupled amplifier.

19-8. THERMAL RUNAWAY

Let us check the thermal resistance requirement of the 2N109 transistor used in the design of the amplifier in Section 19-4. This means the thermal stability equation covered in Section 13-9 must be used, which requires the following:

$$[V_{CC} - 2I_{C,Q}(R_e + R_C)](S)(0.07 I_o) < \frac{1}{\theta}. \tag{19-10}$$

For the amplifier in Fig. 19-3, the following values exist:

$V_{CC} = 9$ V, $I_{C,Q} = 15.5$ mA, $R_e = 47$ ohms, R_C (transformer primary dc resistance) = 20 ohms, S = 10 (this was designed into the circuit), $I_o =$ 7 μA.

Substituting these values in the above equation yields:

$$35 \times 10^{-6} < \frac{1}{\theta}, \text{ or } \theta < 28 \times 10^3 \text{ °C/W}.$$

The 2N109 (assuming its physical characteristics are very similar to the 2N104) has a junction-to-ambient thermal resistance $\theta_{J\text{-}A}$ of 400°C/W.

Thus, this transistor easily meets the requirements set out by the equation, and the circuit is in no danger of thermal runaway. However, this condition is not so easily met in large power amplifiers, where it is not unusual to need a transistor with a $\theta_{J\text{-}A}$ less than 3°C/W.

19-9. OUTPUT TRANSFORMER SATURATION

In a single-ended transformer-coupled amplifier, the relatively large dc collector current flowing under zero signal conditions may cause a large magnetic field density in the transformer core. Then, when the collector current (and magnetizing force) is made to increase due to an increase in input signal, it is possible that there may not be a correct corresponding increase in magnetic flux to induce the proper voltage on the secondary.

On the negative swing, however, the collector current would be able to decrease and reduce the flux by the correct amount, inducing in the secondary a voltage larger than was obtained on the positive half-cycle, as illustrated in Fig. 19-6.

19-10. DISADVANTAGES OF A SINGLE-ENDED TRANSFORMER-COUPLED AMPLIFIER

(1) The total harmonic distortion is unacceptably high for most audio applications except in cheaper amplifiers. The second harmonic distortion is usually the greatest offender.

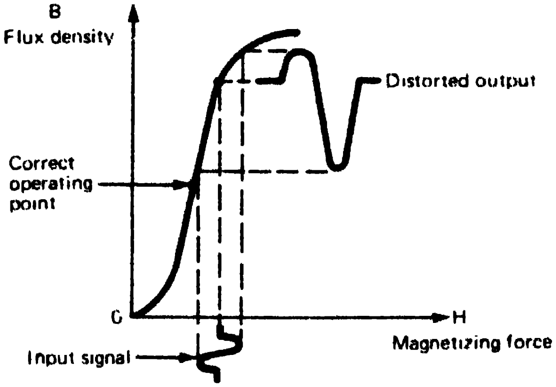


Fig. 19-6. Graph showing how transformer saturation can produce a distorted output.

(2) Maximum output power is developed at a load resistance that is different from the value for minimum harmonic distortion. Consequently, the amplifier is designed to work into a load that minimizes distortion with a resulting sacrifice in output power.

(3) The output transformer is subject to saturation problems due to the dc current in the primary.

It should not be inferred from the above that the single-ended transformer-coupled amplifier does not have its place among amplifiers, although it is seldom used where high power output is required. With suitable negative feedback and proper transformer selection, it is often used as a driver stage for subsequent amplifiers, as shown in the next chapter.

19-11. PUSH-PULL AMPLIFIERS

If in some way the largest contributor to the total harmonic distortion (the second harmonic) could be eliminated, then the load resistance could be adjusted to work nearer to the maximum output power. A method to accomplish this second harmonic cancellation is found in the push-pull circuit, which employs two active devices, requires two input signals 180° out of phase with each other, and usually needs a center-tapped output transformer. It will be shown that this type of circuit effectively eliminates all of the three disadvantages listed in Section 19-10.

Fig. 19-7 is a circuit of a class A push-pull amplifier. By class A, we mean that current flows in the output of the active device (each transistor) for the whole of the input cycle. It is never cut off, as it is in a class B amplifier, which is considered in Chapter 20.

19-12. DESCRIPTION OF OPERATION OF A CLASS A PUSH-PULL AMPLIFIER

The input or driver transformer, T1, provides a phase-splitting function in which two voltages 180° out of phase with each other are produced at the bases of the transistors. See Fig. 19-8. (Note the polarity dots on T1 secondary.)

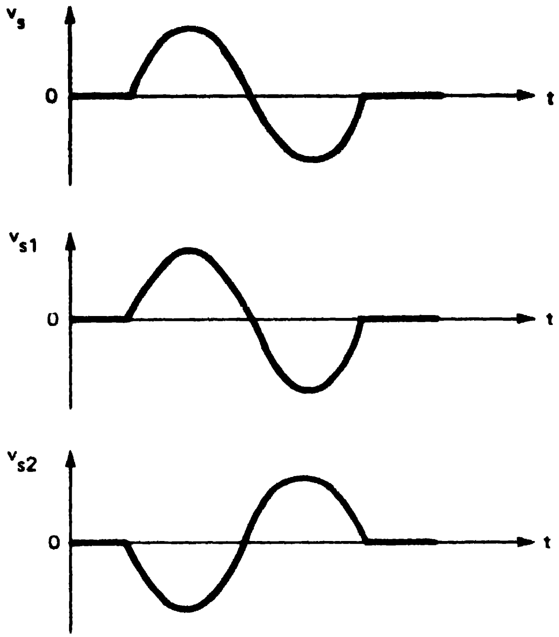


Fig. 19-8. Phase relationships of voltages at the output of the driver transformer, T1, compared with the input, v_s

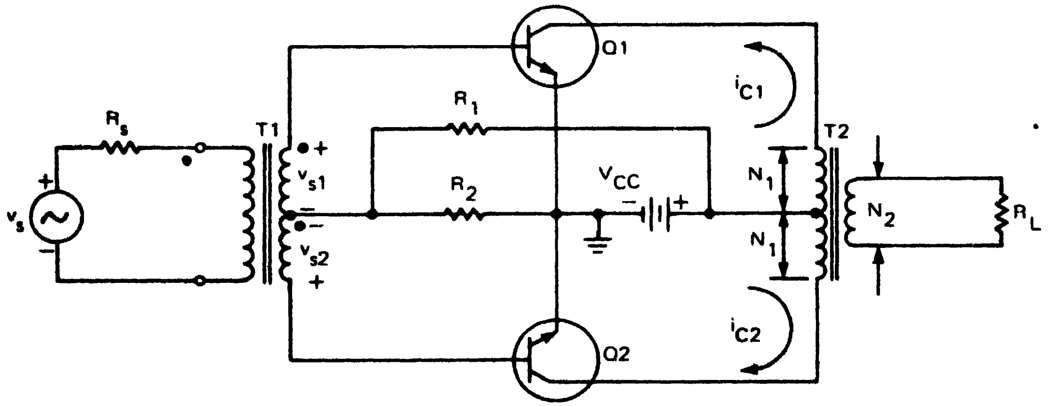


Fig. 19-7. Circuit diagram of a class A push-pull amplifier.

The ratios of v_{s1} and v_{s2} to v_s will depend of course upon the turns ratio of T1. The transformer is usually chosen to provide proper matching between R_s and the input resistance to the transistor bases, so that the secondary voltages may be larger or smaller than v_s, depending upon the circuit application.

Resistors R_1 and R_2 are chosen to provide the correct amount of forward bias (positive voltage in this case) on the transistor bases, so that collector current will flow in each transistor for the whole cycle of input signal.

In the output circuit, V_{CC} is tied to the transistor collectors through the center-tapped output transformer, T2. In this class A push-pull circuit, the dc forward bias on the transistors (which must have matched characteristics for proper operation) will provide a dc collector current I_{C1} and I_{C2} in the transistors. Since these currents pass through the windings of T1 in opposite directions, their magnetizing effects will cancel, with the result that there are no magnetic saturation problems in a push-pull output transformer.

Consider now v_s going positive. Using the polarity dots on T1, this will mean that the base of Q1 will become more positive, while the base of Q2 will become less positive. Evidently, i_{C1} will increase, and i_{C2} will decrease. The voltage that will be induced in the secondary of T2 across R_L will be proportional to the *difference* in the collector currents, $i_{C1} - i_{C2}$, since their magnetizing forces are opposite.

Similarly, when v_s goes negative, the base of Q1 will become less positive (base current i_{B1} will decrease with a corresponding decrease in i_{C1}), and the base of Q2 will become more positive (increasing i_{B2} and i_{C2}). The voltage across R_L is again due to the difference between i_{C1} and i_{C2}, but since i_{C2} is now greater than i_{C1} the polarity of voltage induced across R_L is reversed. This "pushing and pulling" in the output circuit is responsible for decreasing the harmonic distortion by the cancelling of all the *even* harmonics, without increasing the odd harmonics, as shown in the following section.

19-13. THEORY OF OPERATION OF A CLASS A PUSH-PULL AMPLIFIER

Assume the base input currents are sinusoidal in shape, but, to simplify the expression for the output current, let $i_{b1} = I_{bm} \cos \omega t$ and $i_{b2} = -I_{bm} \cos \omega t$, where the minus sign indicates the 180° phase displacement.

The collector current of Q1 is then given (as in the section on harmonic distortion, 7-12) by

$$i_{C1} = I_{C,Q} + A_0 + A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + A_4 \cos 4\omega t. \qquad (19\text{-}11)$$

$$\text{Since } i_{b2} = -i_{b1} = I_{bm} \cos(\omega t + \pi), \qquad (19\text{-}12)$$

the collector current for Q2 is found by replacing ωt by $(\omega t + \pi)$ in the equation for i_{C1}.

Thus,

$$i_{C2} = I_{C,Q} + A_0 + A_1 \cos(\omega t + \pi) + A_2 \cos 2(\omega t + \pi) + A_3 \cos 3(\omega t + \pi) + A_4 \cos 4(\omega t + \pi), \qquad (19\text{-}13)$$

which yields

$$i_{C2} = I_{C,Q} + A_0 - A_1 \cos \omega t + A_2 \cos 2\omega t - A_3 \cos 3\omega t + A_4 \cos 4\omega t. \qquad (19\text{-}14)$$

(Note the alternating signs making the odd frequency components negative.)

But the output voltage is dependent upon $i_{C1} - i_{C2}$:

$$\therefore i_{C1} - i_{C2} = (I_{C,Q} + A_0 + A_1 \cos \omega t + A_2 \cos 2\omega t + A_3 \cos 3\omega t + \ldots)$$

$$- (I_{C,Q} + A_0 - A_1 \cos \omega t$$

$$+ A_2 \cos 2\omega t - A_3 \cos 3\omega t + \ldots).$$

$$\therefore i_{C1} - i_{C2} = 2A_1 \cos \omega t + 2A_3 \cos 3\omega t$$

$$+ 2A_5 \cos 5\omega t + \ldots . \qquad (19\text{-}15)$$

Note that the dc components and *all* the even harmonics have cancelled, leaving the third, $2A_3$, as the main source of distortion now, instead of the second.

Since the percent third harmonic distortion is given by

$$D_3 = \frac{2A_3}{2A_1} \times 100\% = \frac{A_3}{A_1} \times 100\%, \qquad (19\text{-}16)$$

and is the same as the third harmonic distortion in a single-ended amplifier, we can expect a great decrease in total harmonic distortion, now given by

$$D_n \qquad D_3{}^2 + D_5{}^2 \qquad (19\text{-}17)$$

The above discussion assumes two identically matched transistors. Deviation from this assumption will yield the presence of even harmonics, since perfect cancelling will not result.

This cancelling effect will also reduce any ripple still present in the power supply, V_{CC}, after filtering. However, the power supply hum will not be eliminated completely, since in a multistage amplifier hum will be present on the earlier voltage amplifier stages and will compose some of the input to the two transistor bases in the push-pull circuit.

19-14. SUMMARY

(1) In a class A amplifier utilizing a direct-coupled resistive load, the maximum conversion efficiency, $\eta = \frac{P_L}{P_B}$, is 25%.

(2) Maximum transistor dissipation occurs at quiescent conditions in a class A amplifier, and the maximum ac power the transistor can deliver, P_L, is equal to $\frac{P_T}{2}$, where P_T is the transistor's maximum dissipation capability.

(3) When a signal is applied to a class A amplifier, the power dissipated in the transistor decreases by an amount equal to the ac power delivered to the load.

(4) A transformer-coupled class A amplifier has a maximum theoretical conversion efficiency of 50%, but its maximum ac power delivering capability is still one-half of its maximum dissipation capability.

(5) The operating point of a transformer-coupled amplifier is located at

$$I_C = \frac{2P_L{}'}{V_{CC}} \text{ and } V_C = V_{CC}$$

if emitter and collector resistances are ignored. $P_L{}'$ is the power delivered to the primary of the output transformer.

(6) The design of a transformer-coupled power amplifier with an emitter-stabilizing resistor requires the selection of a stability factor S to provide a closed solution.

(7) Calculation of the total harmonic distortion is accomplished using the ac load line on the characteristic curves and the equations of Section 7-13.

(8) Calculation of output power should be done using the fundamental;

$$P_1 = \frac{A_1{}^2 R_L{}'}{2}$$

since distortion may invalidate the use of

$$P = V_{c,p\text{-}p} \times I_{c,p\text{-}p}$$

(9) A graph of power output and total harmonic distortion against load resistance shows maximum output power occurring at a higher value of load resistance than that which provides minimum distortion.

(10) A single-ended transformer-coupled amplifier has two additional disadvantages to the above: The total harmonic distortion is unacceptably high with the second harmonic contributing the most; dc current in the transformer primary causes saturation problems.

(11) A class A push-pull amplifier consists of two active devices, driven by two signals 180° out of phase with each other, and a center-tapped output transformer.

(12) The operation of a push-pull amplifier is such as to cause a cancelling of all the even harmonics without a change in the fundamental and odd harmonics, so distortion is greatly reduced.

QUESTIONS

Q19-1. Explain how the active device (transistor) in a class A direct-coupled-resistive-load amplifier dissipates less power when a signal is applied than with no signal. Is this also true in a transformer-coupled-load power amplifier?

Q19-2. What principal factor contributes to the doubling of the conversion efficiency in a transformer-coupled-load amplifier compared with an amplifier with a direct-coupled resistive load? What other advantages does the transformer provide?

Q19-3. (a) Under what conditions would the dc load line in a transformer-coupled amplifier *not* be drawn vertically on the collector characteristics? Which way would it slope?

(b) What are the equations that specify the operating point in terms of the output power delivered to the primary?

(c) What should be the effective ac resistance reflected into the primary?

Q19-4. In designing a power amplifier and selecting the bias resistors, what additional decisions must be made to provide a closed solution?

Q19-5. Describe how you would calculate the total harmonic distortion of a CE amplifier with a transformer-coupled load driven from a source of low internal resistance. Explain what curves you would need, etc.

Q19-6. What two equations may be used to calculate the output power of an amplifier? When will they give the same result?

Q19-7. Why may an amplifier be operated at a load resistance lower than the value that would provide maximum output power?

Q19-8. What disadvantages are there in a single-ended transformer-coupled power amplifier? Explain transformer saturation.

Q19-9. (a) What is required to provide push-pull operation?

(b) How is transformer saturation avoided in this type of operation?

(c) What other advantages does push-pull provide?

Q19-10. Draw in proper phase relationship the waveforms of v_s, v_{s1}, v_{s2}, i_{b1}, i_{b2}, i_{C1} and i_{C2} for the amplifier in Fig. 19-7.

PROBLEMS

P19-1. A transistor's maximum average dissipation capability, P_T, is 2.5 W. When used in a class A direct-coupled power amplifier determine:

(a) The maximum ac power deliverable to the load.

(b) The minimum dc power needed from the power supply to provide the power in (a).

(c) The dc power dissipated in the load and in the transistor, with no input signal, for the conditions in (b).
(d) The dc power in the load and transistor for the conditions in (a).

P19-2. Derive Eq. 19-1 in the following manner: Assume an ac load line with a peak current of I_C and $V_{CE} = 0$ (saturation) and a peak voltage of V_C and $I_C = 0$ (cutoff). Assume the operating point is at the middle of the load line, and the signal drives the transistor from one extreme of the ac load line to the other. Obtain P_D and P_L.

P19-3. Repeat P19-1 assuming a transformer-coupled load.

P19-4. Design a power amplifier, as in Fig. 19-3, using a germanium 2N176 transistor, which has a P_T of 10 W. The power required at the load is 3 W. Assume an output transformer efficiency of 75%, $R_e = 5$ ohms, $V_{CC} = 20$ V, and a base-emitter bias voltage of 0.4 V. Use a stability factor, S, of 10. Specify:
(a) The primary impedance of the output transformer.
(b) The operating point (approximate).
(c) The values of R_1 and R_2.

P19-5. Design a power amplifier, as in Fig. 19-3, to provide 250 mW to a 100-ohm load, R_L. The transistor is a silicon 40309 with $P_T = 1$ W at ambient temperatures up to 25°C. Assume a transformer efficiency of 60%, $R_e = 22$ ohms, $S = 20$, $V_{CC} = 12$ V, and a base-emitter bias voltage of 0.8 V. Specify:
(a) The primary impedance of the output transformer.
(b) The operating point (approximate).
(c) The values of R_1 and R_2.

P19-6. (a) Check the amplifier in P19-4 to see if it is stable against thermal runaway, given resistance of transformer primary = 10 ohms, $I_{CBO} = I_o = 3$ mA, $\theta_{J\text{-}A} = 10$°C/W.
(b) What is the maximum value of S that could be tolerated before thermal runaway would be a problem?

P19-7. Check the amplifier in P19-5 to see if it is stable against thermal runaway, given resistance of transformer primary = 30 ohms, $I_{CBO} = I_o = 0.25\ \mu$A, $\theta_{J\text{-}A} = 175$°C/W.

P19-8. Use the fact that

$$\cos(A + B) = \cos A \cos B - \sin A \sin B,$$

along with $\cos \pi = -1$, $\cos 2\pi = 1$, $\sin \pi = \sin 2\pi = 0$, to show that Eq. 19-13 yields Eq. 19-14.

P19-9. A single transistor working in a class A amplifier with a transformer-coupled load produces harmonic amplitudes in the output of: $A_o = 1.5$ mA, $A_1 = 120$ mA, $A_2 = 10$ mA, $A_3 = 4$ mA, $A_4 = 2$ mA, $A_5 = 1$ mA.
(a) Determine the percent total harmonic distortion.
(b) Assume a second identical transistor is used along with a suitable transformer to provide push-pull operation. Use the above amplitudes to determine the new total harmonic distortion in the output.

20

TRANSISTOR POWER AMPLIFIERS II: CLASS B

20-0. INTRODUCTION

The class A push-pull amplifier was seen to remove three drawbacks found in the single-ended transformer-coupled amplifier. Most importantly, it reduced harmonic distortion significantly. However, the conversion efficiency is still only 50%. A method of increasing this efficiency as well as the output power for a given transistor dissipation capability is found in the class B push-pull amplifier. In this chapter, we shall design a class B amplifier and examine ways to dispense with the transformers as, for example, in the complementary symmetry amplifier.

20-1. THE CLASS B PUSH-PULL AMPLIFIER

This amplifier, like the class A amplifier, requires a pair of active devices, but in this case each transistor "looks after" one-half of the signal. Both transistors are biased at cutoff. This is easily done by providing zero bias on the base of each transistor by connecting the emitter and base together (as far as dc is concerned). This is illustrated in Fig. 20-1 where, for circuit variation, PNP transistors have been used.

As before, T1 is a phase splitter providing two signal voltages 180° out of phase. With no signal applied, both Q1 and Q2 are cut off, and there is no current drain on the supply V_{CC} except for the reverse leakage current, I_{CEO}. Thus, there is no power wasted on standby, and the dissipation in the transistors is practically zero.

Now assume v_s goes positive. Since this will make the base of Q1 positive, Q1 will be reverse biased and will not conduct any current, but Q2 will be forward biased, allowing i_{C2} to flow through T2. When v_s goes negative, Q2 will be cut off, and Q1 will "look after" the negative half cycle of the input signal. Therefore, only one transistor conducts at a time, each one handling one-half of the input signal. The output transformer serves to "join together" these two halves, producing in R_L a full sine wave. The output voltage is again dependent upon the difference of the collector currents, so that even harmonics will still cancel.

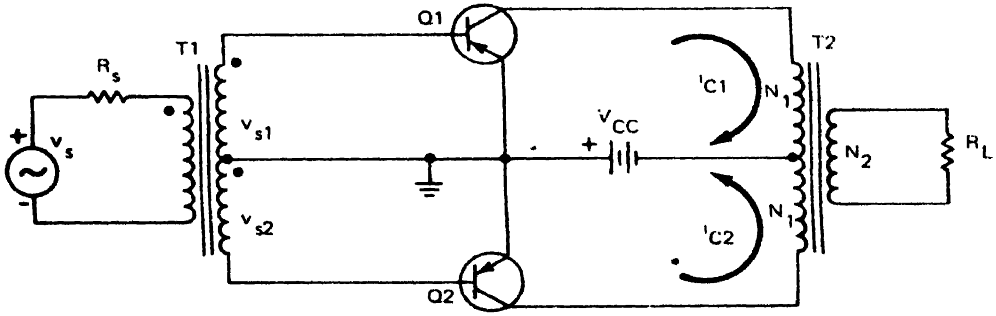


Fig. 20-1. Circuit of a class B push-pull amplifier.

However, since the joining is not perfect (as will be described), the harmonic distortion is higher than in a class A push-pull amplifier. The diagrams in Fig. 20-2 illustrate the class B action.

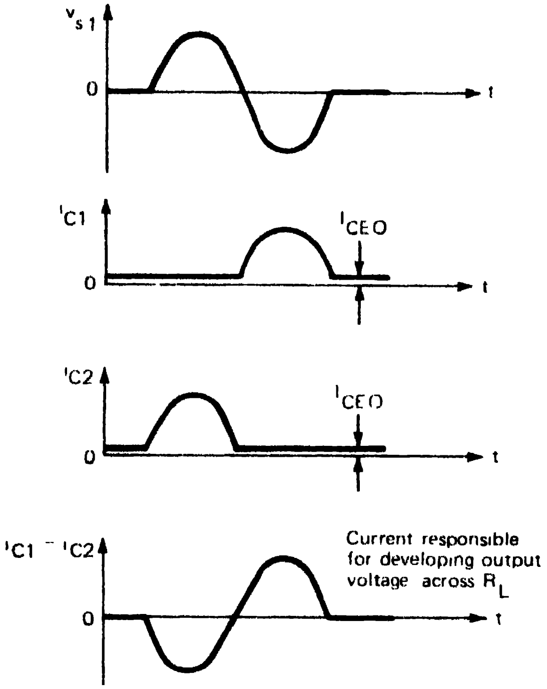

vs1
0
t
IC1
ICEO
0
t
IC2
ICEO
0
t
IC1 − IC2
Current responsible for developing output voltage across RL
0
t

Fig. 20-2. Diagrams illustrating class B action.

It can be seen that the collector current increases from zero to some peak value quite rapidly, so that the power supply must have a good regulation.

20-2. CROSSOVER DISTORTION

The above diagrams were drawn with the assumption that each transistor's collector current was a perfect half sine-wave, so that the joining together was possible without any misalignment. In actual practice, true class B operation introduces *crossover distortion*. This arises from the fact that at low collector currents (near cutoff) the value of β is quite low compared with the value for the rest of the cycle, so if the input base voltage is sinusoidal the output collector current will not be a faithfully enlarged sine wave at low values of current. The result is shown in Fig. 20-3.

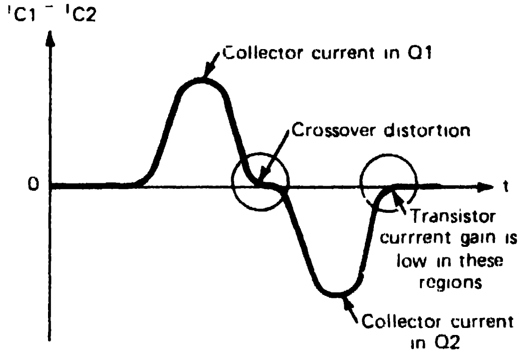

IC1 − IC2
Collector current in Q1
Crossover distortion
0
t
Transistor currrent gain is low in these regions
Collector current in Q2

Fig. 20-3. Diagram showing how crossover distortion occurs.

That is, the transistor remains cut off until the base voltage exceeds 0.2 V for germanium and 0.5 V for silicon. The resulting crossover distortion is objectionable to the ear as can be demonstrated in the laboratory.

20-3. CLASS AB PUSH-PULL AMPLIFIER

To avoid this effect, the base-emitter junctions of both transistors are slightly forward biased by the amounts mentioned above, so that when an ac signal is applied to the base, collector current starts flowing immediately. This biases the transistor away from class B (slightly) toward class A; the resulting designation is class AB. This is achieved by adding biasing resistors as shown in Fig. 20-4. Emitter stabilizing resistors are also shown. At first sight this looks like a class A push-pull circuit, but

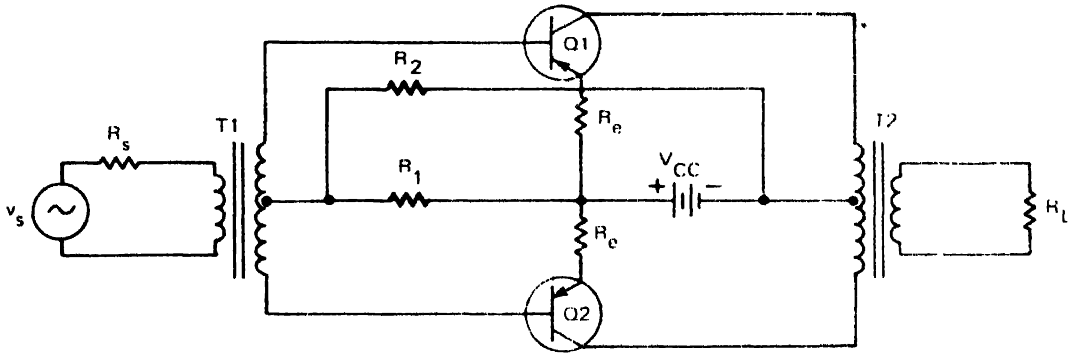


Fig. 20-4. Circuit of a class AB push-pull amplifier, showing the biasing resistors R_1 and R_2.

the choice of R_1 and R_2 places the operating point much closer to cutoff and class B operation. The insertion of R_e will cause some decrease in output power due to the degenerative feedback effect, but this disadvantage is more than offset by the decrease in distortion.

20-4. TRANSISTOR PHASE INVERTER

In place of using a center-tapped input transformer to supply two signals 180° out of phase, a transistor phase inverter may be used. This is shown in Fig. 20-5.

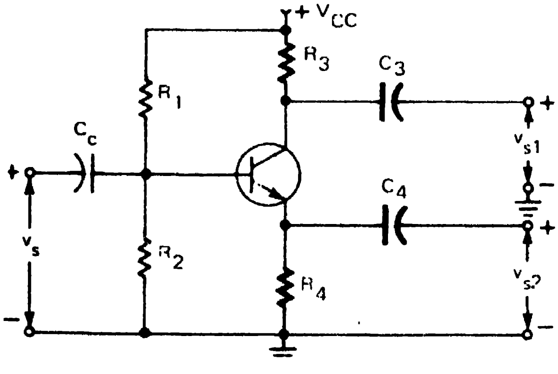


Fig. 20-5. Transistor phase inverter to provide two signals 180° out of phase with each other.

R_3 and R_4 are chosen equal to each other, as are C_3 and C_4. Signal v_{s2} is in phase with v_s (emitter follower), while v_{s1} is 180° out of phase with v_s. There are three drawbacks associated with this circuit:

(1) Since this is of the emitter-follower type, there is no voltage gain.

(2) Since the emitter and collector currents are not equal, the output voltages v_{s1} and v_{s2} will have some imbalance.

(3) This phase inverter could not be used to drive a class B push-pull amplifier, since v_{s1} and v_{s2} are only equal if driving a resistive load over the *full* cycle. It could be used on a class A push-pull circuit, where the load is constant on each output.

20-5. CONVERSION EFFICIENCY OF A CLASS B AMPLIFIER

To indicate the operating point and ac load line for two transistors in push-pull, it is necessary to draw the composite characteristic curves for two transistors. These are shown in Fig. 20-6.

Let I_c be the rms value of the output collector current in the primary of the output

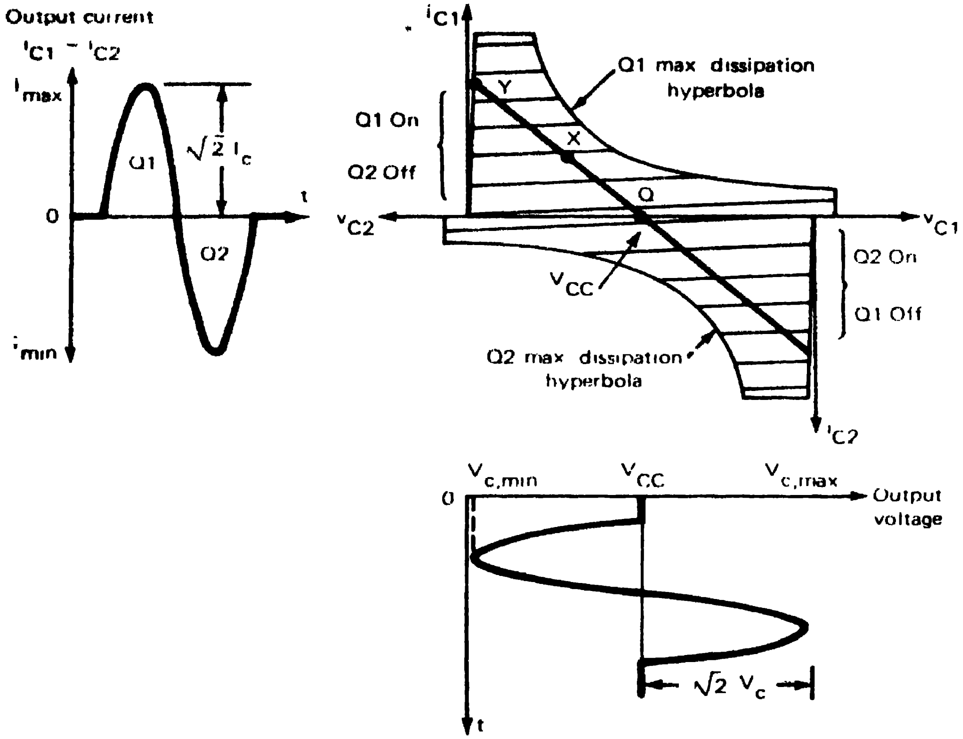

Output current
$i_{C1} - i_{C2}$
I_{max}
Q1
$\sqrt{2}\,I_c$
0
t
Q2
i_{min}
i_{C1}
Q1 max dissipation hyperbola
Y
X
Q
Q1 On
Q2 Off
v_{C2}
v_{C1}
Q2 On
Q1 Off
V_{CC}
Q2 max dissipation hyperbola
i_{C2}
$V_{c,min}$
V_{CC}
$V_{c,max}$
0
Output voltage
$\sqrt{2}\,V_c$
t

Fig. 20-6. The composite characteristic curves of the two transistors, showing the ac load line and output current and voltage.

transformer, given by $i_{C1} - i_{C2}$, and let V_c be the rms value of the output collector-to-collector voltage on the transformer primary. Then the ac power delivered to the reflected resistance R_L' is given by $P_L' = V_cI_c$.

The dc power supplied by the battery is given by $P_B = V_{CC}I_C$, where I_C is the average value of $i_{C1} + i_{C2}$, a full-wave-rectified sine wave, as shown in the partial circuit of Fig. 20-7.

The peak value of the current delivered by the battery is $\sqrt{2}I_c$ (see Fig. 20-6). Recalling that the average value of a full-wave-rectified sinusoidal waveform is given by

$$I_{avg} = \frac{2I_m}{\pi},$$

the average current delivered by V_{CC} is

$$I_C = I_{avg} = \frac{2\sqrt{2}I_c}{\pi}.$$

Assuming $V_{c,min} \approx 0$, then $V_{CC} = \sqrt{2}V_c$

$$\therefore P_B = V_{CC}I_C = \sqrt{2}V_c \times \frac{2\sqrt{2}I_c}{\pi} = \frac{4V_cI_c}{\pi}$$

But $V_cI_c = P_L'$.

$$\therefore P_B = \frac{4}{\pi}P_L'.$$

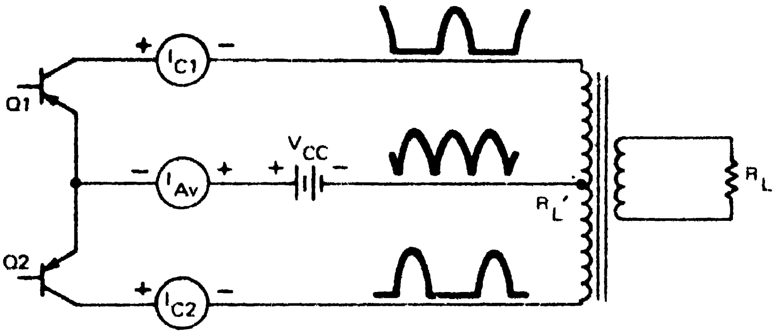


Fig. 20-7. Partial circuit of a class B push-pull amplifier showing the current delivered by the dc supply.

Conversion efficiency, η × 100%

$$= \frac{\pi}{4} \times 100\% = 78.5\%. \qquad (20\text{-}1)$$

This is the optimum conversion efficiency possible in a pure class B push-pull circuit. If class AB is used, the efficiency will drop, approaching 50% as the bias point approaches class A operation. This is simulated in the composite curves by "overlapping" the two characteristics. Usually, class AB operation (with bias stabilizing resistors, R_e) typically reduces the efficiency by approximately 8 to 10%. The resulting 70% efficiency is typical of class AB operation and is welcome in high-powered amplifiers.

20-6. RELATION BETWEEN MAXIMUM OUTPUT POWER AND LOAD RESISTANCE

The resistance, R_L', seen by each transistor (since only one transistor conducts at a time) is actually the resistance reflected from the load into *one-half* of the primary of the transformer and is given by the following.

Referring to Fig. 20-6, the slope of the ac load line is, by an earlier discussion, equal to $\frac{1}{R_L'}$, and is equal to $\frac{\sqrt{2}I_c}{V_{CC}}$, assuming $V_{c,min} \approx 0$.

Thus, $R_L' = \frac{V_{CC}}{\sqrt{2}I_c}$.

Now $P_L' = I_c V_c$.

$\therefore I_c = \frac{P_L'}{V_c}$, and, since $V_{CC} = \sqrt{2}V_c$,

$$I_c = \frac{P_L'\sqrt{2}}{V_{CC}}, \text{ so that}$$

$$R_L' \quad \frac{V_{CC}}{\sqrt{2}} \times \frac{V_{CC}}{\sqrt{2}P_L'} = \frac{V_{CC}^2}{\quad} \qquad (20\text{-}2)$$

where P_L' is the total ac power delivered to the transformer and is realized only when the input signal provides a full swing over the load line.

Defining α to be the turns ratio of one-half of the primary to the secondary, $\frac{N_1}{N_2}$ (see Fig. 20-1), then $R_L' = \alpha^2 R_L$. But the output transformer is selected on the basis of the collector-to-collector resistance, $R_{C\text{-}C}$, which now doubles the turns ratio from the total primary to the secondary.

Thus, $R_{C\text{-}C} = (2a)^2R_L = 4a^2R_L = 4R_L'$. (20-3)

R_L' is thus determined by V_{CC} and P_L', and in turn determines the collector-to-collector resistance of the output transformer. The question is, what is the limit on P_L' in a class B push-pull amplifier? Evidently, it will have something to do with the dissipation capability of the transistors being used.

20-7. RELATION BETWEEN MAXIMUM OUTPUT POWER AND TRANSISTOR DISSIPATION

Refer to the ac load line drawn on the composite curves for the two transistors in push-pull considered earlier in Fig. 20-6. The way in which the *instantaneous* dissipation varies in the transistor as the operating point moves along the load line from Q to Y and back to Q is shown graphically in Fig. 20-8.

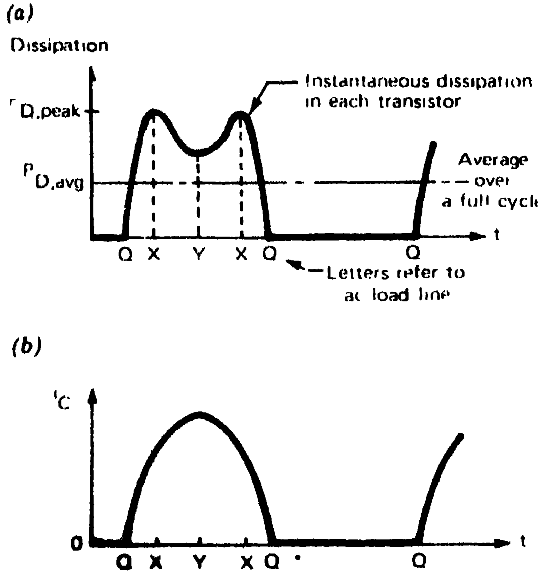


Fig. 20-8. (a) Diagram showing how the instantaneous dissipation in each transistor varies with the collector current shown in (b).

It is seen that the transistor passes through some peak dissipation, $P_{D,peak}$, very rapidly in the region of X. If this peak dissipation capability is given, then the maximum permissible total power deliverable to the transformer without exceeding $P_{D,peak}$ is given by

$$P_L' = 2P_{D,peak}. \qquad (20\text{-}4)$$

However, it is seen from the above diagram that the transistor is cut off for one-half of a cycle, allowing it to cool so that some *average* power is dissipated in the transistor over a full cycle that is lower than the peak instantaneous dissipation. Actually, what the transistor manual specifies is the *average* dissipation capability, P_T, defined as the difference between the dc input power to the collector and the ac output power to the load. Using this definition, it can be shown that the maximum (worst case) average dissipation *per transistor*, $P_{D,max} = 0.2P_L'$ or,

$$P_L' = 5P_D \qquad (20\text{-}5)$$

where P_L' is the total maximum power obtainable (from two transistors in class B push-pull) if the base current varies over the full ac load line given by R_L' obtained earlier.* Table 20-1 summarizes the equations and provides a comparison of power requirements for an ac output power of 10 watts.

---

*Normally this will cause too much distortion in the output, so that the amplifier is usually designed to make $P_L' \approx 0.5$ to $2P_T$, depending upon the transistor's linearity and acceptable level of distortion.

TABLE 20-1 COMPARISON OF POWER REQUIREMENTS IN THREE-TYPES OF TRANSISTOR AMPLIFIERS

	CLASS A SINGLE TRANSISTOR DIRECT-COUPLED	CLASS A SINGLE TRANSISTOR TRANSFORMER COUPLED	CLASS B TWO TRANSISTORS PUSH-PULL
Ac Output Power	$P_L = 10$ W	$P_L' = 10$ W	$P_L' = 10$ W
Transistor Dissipation	$P_T = 2P_L = 20$ W	$P_T = 2P_L' = 20$ W	$P_T = 0.2P_L' = 2$ W (per transistor)
Conversion Efficiency	$\eta = 25\%$	$\eta = 50\%$	$\eta = 78.5\%$
Dc Power Input	$P_B = 40$ W	$P_B = 20$ W	$P_B = 12.75$ W
Transformer Primary Resistance (Reflected)		$R_L' = \frac{V_{CC}^2}{2P_L'}$	$R_L' = \frac{V_{CC}^2}{2P_L'}$ $R_{C\text{-}C} = 4R_L'$

20-8. DESIGN OF A CLASS B PUSH-PULL AMPLIFIER

Example 20-1. To illustrate the theory in the previous section, let us design a class B push-pull circuit to deliver 200 mW to a 4-ohm load using 2N109 transistors. It is required to select the values of the components in Fig. 20-9 including the transformers.

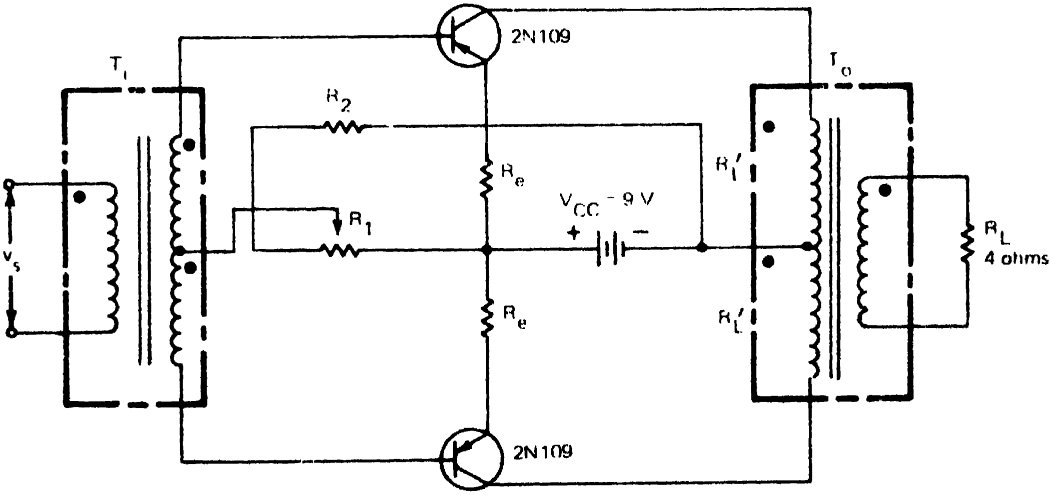


Fig. 20-9. Circuit of a class B amplifier to be designed for a power output of 200 mW in R_L.

Solution. If $P_L = 200$ mW, then, allowing for a 70% efficient output transformer, T_o, the power delivered to the primary of the transformer is

$$\frac{200}{0.7} \quad 285 \text{ mW}.$$

$V_{CE,max}$ for 2N109 = 25 V; therefore V_{CC} must be less than 25/2 = $12\frac{1}{2}$ V. Let us use $V_{CC} = 9$ V, since this is the nearest standard battery size.

Now check to see if the 2N109 can dissipate enough power.

$P_{D,max(avg)} \approx 0.2P_L' = 0.2 \times 285 = 57$ mW, which is well within the 165 mW average rating of the 2N109 at 25°C.

Now select transformer T_o. $R_L \quad \frac{V_{CC}}{2P_L'}$.

$$\therefore R_L' = \frac{-}{2 \times 285 \times 10^{-3}} = 142 \text{ ohms}.$$

But the collector-to-collector resistance, $R_{C\text{-}C} = 4R_L' = 568$ ohms. Thus, T_o should be a center-tapped transformer, 600-ohm primary, 4-ohm secondary, with a minimum power rating of 0.3 W.

Assuming the transistor is driven over the extreme of the ac load line, the peak collector current in each transistor is

$$I_{c,peak} \quad \frac{V_{CC}}{R_r'} \quad \frac{9}{142} = 63 \text{ mA}.$$

The dc component of this collector (and emitter) current is $\frac{63}{\pi} = 20$ mA, which causes a dc voltage drop across the bias stabilization resistor, R_e, (assumed to be 10 ohm) of 0.2 V. However, with no signal applied, the collector idling current will be practically zero, and the voltage developed across R_e that tends to reverse bias the base-emitter junction of the transistors will be negligible. Let it be required that the voltage divider consisting of R_1 and R_2 provide a forward bias of 0.2 V (0.6 V for a silicon transistor). R_1 has been indicated as variable, so that it may be adjusted for class AB operation.

Using the voltage divider equation,

$$0.2 \text{ V} = 9 \text{ V} \times \frac{R_1}{R_1 + R_2}$$

from which $R_2 = 44R_1$.

Thus, if $R_1 = 10$ ohms, $R_2 = 440$ ohms (470 ohms), or if $R_1 = 100$ ohms, $R_2 = 4.4$ K ohms (4.7 K ohms). The higher resistance will minimize the drain on V_{CC}, although the low input resistance to the transistor will affect the bias resistor values.

To select the input or driver transformer, T_i, it is necessary to know the input impedance. Since R_e (10 ohms) is unbypassed, $R_i \approx h_{ie} + h_{fe}R_e = 0.1 \text{ K} + 100 \times 10 = 1.1$ K ohms (h_{ie} is quite small when the transistor is operated with large signals). This is from base to the center tap. Thus, the resistance from base-to-base is in the order of 4.4 K ohms. Assuming the source v_s has an impedance of 600 ohms (which is typical), select the nearest standard transformer having a 500-ohm primary and a 5-K ohm center-tapped secondary.

Note the conversion efficiency of the circuit:

$$= V_{CC} \times I_C = 9 \times (2 \times 20) \text{ mW} = 360 \text{ mW},$$

$$P_r' = 285 \text{ mW},$$

$$\frac{285}{360} \times 100\% \approx 78.5\%.$$

This is to be expected, since our equations were based on idealized optimum conditions. In practice, the efficiency will be closer to 65 or 70%.

20-9. OTHER CLASS B PUSH-PULL AMPLIFIERS

Due to the expense and weight of transformers, any method that could avoid their use would be welcomed. A circuit that dispenses with the output transformer is shown in Fig. 20-10.

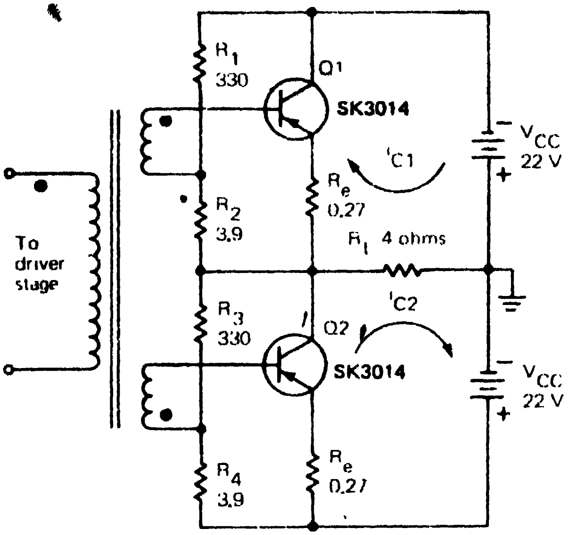


Fig. 20-10. Circuit of a class B push-pull amplifier that does not require an output transformer. (Courtesy of the RCA Corporation.) (All resistance values in ohms.)

Why is it that no transformer is required at the output? It is because high-powered transistors now exist that have a collector-to-emitter output impedance in the order of 4 to 8 ohms. This allows the "single-ended push-pull" operation, in which the collector current of each transistor flows through the load (such as the voice coil of a speaker) directly. The voltage developed across the load is again due to the difference in collector currents, $i_{C1} - i_{C2}$, so this is a true push-pull application.

The total harmonic distortion of this amplifier is 5% with an rms power output of 25 watts. (The EIA, Electronics Industry Association, gives this amplifier a music power output rating of 45 watts, while the IHFM, the Institute of High Fidelity Manufacturers, rates this amplifier at 50 watts music power.)

20-10. COMPLEMENTARY SYMMETRY

A method to remove the input driver transformer (in addition to the output transformer) is easily seen in a transistor amplifier. Consider the function of the driver transformer; it is used to provide two signal voltages 180° out of phase. This is necessary to turn one transistor on as the other transistor is being turned off and vice versa

However, if the same signal voltage were applied to the bases of two transistors, one a PNP and the other an NPN, the actions on the two transistors are opposite. On the positive half cycle, the NPN would be forward biased, but the PNP reversed biased, while, on the negative half cycle, the collector current in the PNP would be increased, and that of the NPN would decrease. It should be evident that no electron-tube counterpart exists, and a complementary symmetry amplifier is not possible using tubes. A class B push-pull solid-state amplifier using the above principle is shown in simplified form in Fig. 20-11.

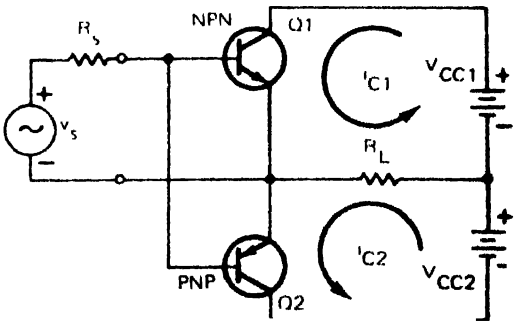


Fig. 20-11. Simplified circuit of a class B push-pull complementary symmetry amplifier that requires no transformers.

The two complementary transistors are acting simultaneously as a phase inverter and an output stage. In order to avoid crossover distortion, each transistor would be slightly forward biased into class AB operation. In the circuit shown, there is complete balance, so that essentially no dc current flows through the load, and direct coupling can be used to a speaker for R_L.

20-11. PRACTICAL COMPLEMENTARY SYMMETRY AMPLIFIER

If the current path is traced out in the basic complementary symmetry amplifier shown earlier, there is a complete series path from the power supplies through the transistors. Moreover, the center of the two power supplies is "floating"—not grounded—, as illustrated in Fig. 20-12.

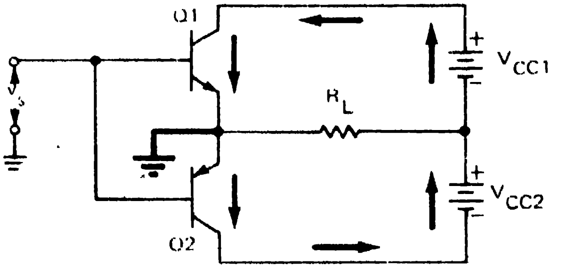


Fig. 20-12. Complementary symmetry amplifier showing nongrounded power supplies.

The ac voltage developed across R_L can be considered to be the ac voltage across either Q1 or Q2.

Although it is not difficult to obtain two voltages of opposite polarity from a power supply, these are usually obtained with respect to ground, whereas in the circuit of Fig. 20-12 the *emitters* are grounded.

A 20-watt complementary symmetry audio power amplifier that utilizes a single polarity power supply (as do most practical complementary amplifiers) is shown in Fig. 20-13. This five-stage amplifier can deliver 20 W of rms power into an 8-ohm load.

Q1 is a common-emitter amplifier whose output is directly coupled to Q2, also a common-emitter amplifier, which in turn supplies the signal to the base of Q3. This is an emitter follower whose output supplies the driver stage Q4. The output of Q4 is connected directly to the bases of the output stages Q5 and Q6.

For proper operation of the output stage, the dc voltage at the center point must stay constant. This is achieved by negative feedback from this point to the base of Q3. The collector load resistance for Q4 is made up of $R_{13} + R_{14}$; C_7 is a "bootstrap" capacitor and is used to keep the current flowing through Q4 constant. Let us analyze the action of the output stage and how the load develops an output. In some complementary symmetry circuits, a diode is added between the bases of Q5 and Q6 as shown in the simplified diagram of Fig. 20-14.

Let us assume that with no signal applied the center-point voltage is 20 V dc (with respect to ground), and the diode has 0.8 V dropped across it (due to R_{13} and R_{14}). Then the base of Q5 is 0.8 V higher than the base of Q6. Under balanced conditions—no applied signal—, the base of Q5 would be 20.4 V, and the base of Q6, 19.6 V. Thus both Q5 and Q6 emitter-base junctions would be forward biased slightly, allowing a small idling current to flow through both transistors. If the ambient temperature should increase (tending to increase current through the transistors), the voltage drop across the diode (also subjected to the increase in temperature) would decrease (say to 0.6 V). Then the forward bias for each transistor would be reduced by 0.1 V and would tend to offset the increase in collector current. Thus, the diode (where

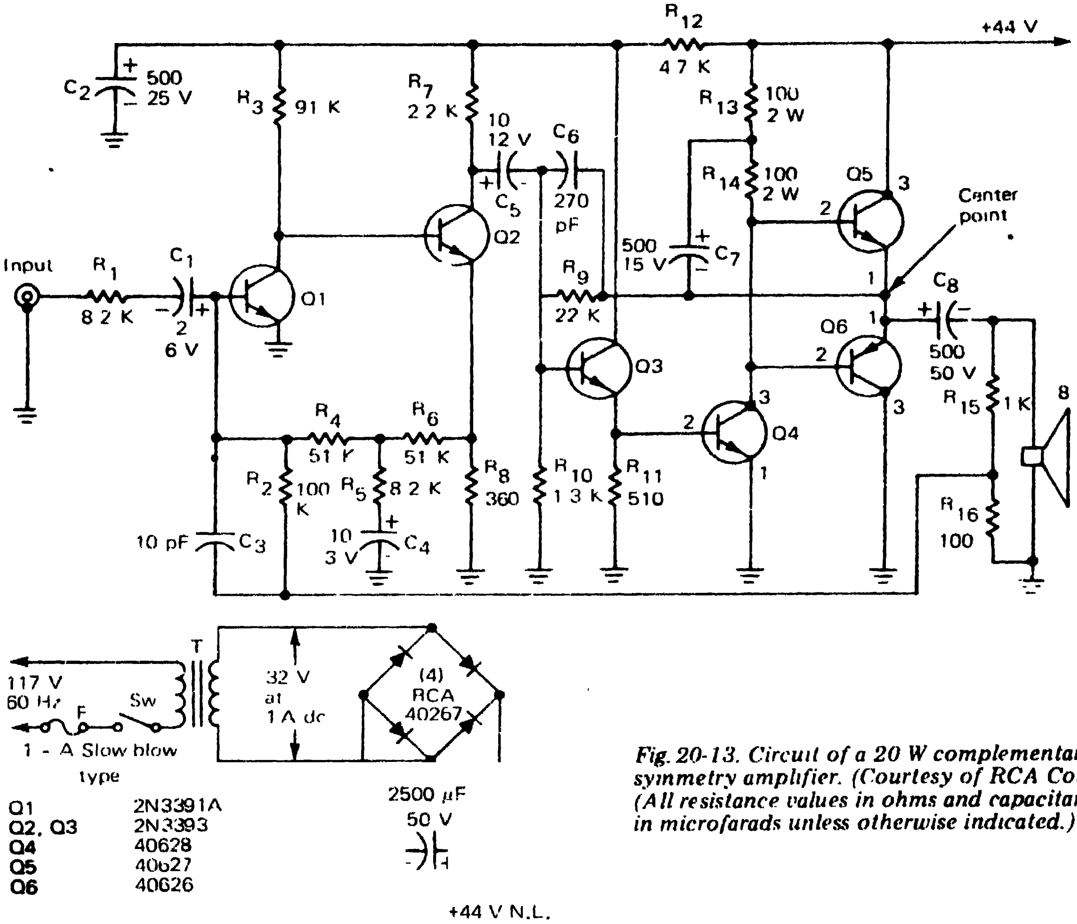

R12
+44 V
C2
500
25 V
R3
91 K
R7
2 2 K
10
12 V
C6
C5
270
pF
47 K
R13
100
2 W
R14
100
2 W
Q5
Center
point
Input
R1
8 2 K
C1
2
6 V
Q1
Q2
500
15 V
C7
R9
22 K
C8
500
50 V
Q6
Q3
R15
1 K
8
R4
51 K
R6
51 K
Q4
R2
100
K
R5
8 2 K
R8
360
R10
1 3 K
R11
510
R16
100
10 pF
C3
10
3 V
C4
T
117 V
60 Hz
Sw
F
1 - A Slow blow
type
32 V
at
1 A dc
(4)
RCA
40267
Q1 2N3391A
Q2, Q3 2N3393
Q4 40628
Q5 40627
Q6 40626
2500 μF
50 V
+44 V N.L.

Fig. 20-13. Circuit of a 20 W complementary symmetry amplifier. (Courtesy of RCA Corporation.) (All resistance values in ohms and capacitance values in microfarads unless otherwise indicated.)

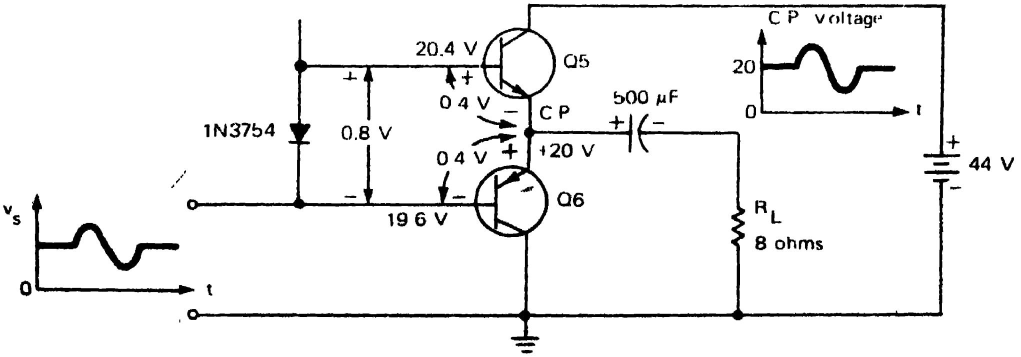

20.4 V
Q5
C P voltage
20
0
t
0 4 V
C P
500 μF
1N3754
0.8 V
0 4 V
+20 V
44 V
Q6
19 6 V
R_L
8 ohms
v_s
0
t

Fig. 20-14. Partial circuit of a complementary symmetry amplifier showing bias voltages.

used) serves as a temperature compensating device, as well as biasing the amplifier into class AB operation.

Now consider the action when v_s goes positive. This will increase the forward bias on Q5, reduce that of Q6. Q5 will conduct more heavily, while Q6 will tend to turn off. Consequently, the voltage at the center point will tend to increase *toward* +44 V during this interval.

On the negative half cycle, Q6 is made to conduct, and Q5 tends to be cut off, so the center point is "dragged down" *toward* ground level or 0 V. (That is, the voltage drop across Q6 from emitter to collector approaches the saturation value of 0.6 V.) The waveform of voltage at the center point is then given by the diagram in Fig. 20-15.

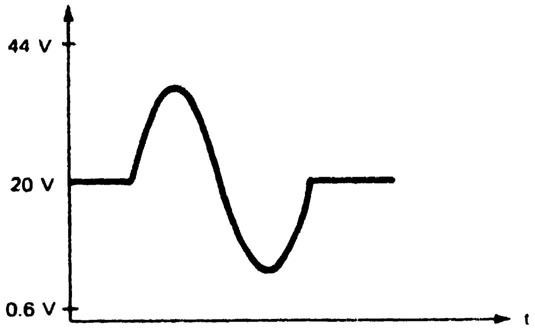


Fig. 20-15. Waveform of the voltage at the center point of Fig. 20-14.

The 500 μF capacitor serves to block the dc component and pass the ac, including even low frequencies.

Typical performance of this amplifier with an 8-ohm load is 1% total harmonic distortion at 1 kHz and 20 watts output, with a low frequency 3-dB point of 20 Hz and flat to above 20 kHz. This is achieved by the large amount of negative feedback used generously throughout the amplifier and the overall feedback from output to input. The hum and noise for this amplifier is 82 dB below 20 watts.

20-12. SUMMARY

(1) A pure class B push-pull amplifier has both of its active devices biased at cutoff, with center-tapped transformers used both at the input and output. Hence, no power is drawn with no signal applied.

(2) Crossover distortion in a pure class B amplifier is the imperfect alignment of the two halves of the output and is due to the fact that the transistors remain cut off when the input is at a low level.

(3) Class AB operation removes the crossover distortion by slightly forward biasing both transistors to permit a small idling current to flow at all times.

(4) The maximum conversion efficiency of a pure class B push-pull amplifier is 78.5%, although a more realistic figure is 70% for class AB operation.

(5) The relation between maximum output power, P_L', and load resistance, R_L', is given by

$$R_L' = \frac{V_{CC}^2}{2P_L'}$$

(6) The output transformer is selected on the basis of the collector-to-collector resistance, $R_{CC} = 4R_L'$.

(7) The transistor dissipation in a class B amplifier is zero under quiescent conditions, but goes through a peak dissipation, $P_{D,peak}$, twice for each cycle of input signal, having an average dissipation over a full cycle of $P_{D,avg}$.

(8) The total power P_L' (from both transistors) deliverable to the output transformer

in a class B amplifier is given by

$$2P_{D,\text{peak}} \quad 5P_T$$

where $P_{D,\text{peak}}$ is the peak dissipation capability of each transistor, and P_T is the maximum average dissipation capability of each transistor.

(9) A summary of power operating conditions for class A and B amplifiers is contained in Table 20-1.

(10) The current flowing from the power supply in a class B amplifier flows in full-wave rectified pulses, so the power supply must have good regulation.

(11) By using high power transistors, it is possible to dispense with the output power transformer by placing the load so that each collector current pulse passes directly through the load, but in opposite directions, thus providing the advantages of push-pull.

(12) Complementary symmetry amplifiers do not depend upon either an input or an output transformer. They utilize opposite types of transistor, PNP and NPN, so that no phase-splitting device is required.

(13) Practical complementary symmetry amplifiers operate from a single-polarity power supply but utilize a large capacitor to couple the load to the center point of the output transistors.

QUESTIONS

Q20-1. (a) What type of bias is used in a true class B push-pull amplifier? How is this obtained in a transistor amplifier?

(b) If each transistor handles one-half of the signal, what component combines the two?

Q20-2. (a) Explain crossover distortion, what it is caused by, and how it is overcome.

(b) What penalty is paid for biasing an amplifier into class AB operation?

Q20-3. (a) Draw the circuit of a PNP transistor phase inverter and indicate the waveforms at the input and the outputs.

(b) What advantage does this circuit have over a transformer phase splitter? What disadvantages does it have?

Q20-4. (a) What assumptions are made in calculating the maximum theoretical efficiency of a class B push-pull amplifier to be 78.5%?

(b) What is the waveform of current flowing from the power supply that feeds a class B push-pull amplifier?

(c) Why should such a power supply have good voltage regulation?

Q20-5. (a) What is the expression that gives the ac resistance reflected into one-half of the primary of the output transformer, in terms of collector supply voltage and output power?

(b) Why is the collector-to-collector resistance seen by the transformer primary *four times* the value mentioned in (a)?

Q20-6. (a) How many times is a transistor subjected to a *peak dissipation* during one cycle of input voltage?

(b) Why is the *average dissipation* of a transistor in a class B push-pull amplifier appreciably lower than the *peak dissipation*?

Q20-7. (a) What is the function of resistors R_1, R_2 and R_3, R_4 in Fig. 20-10?

(b) If all the components are perfectly matched what current flows through R_L when the input signal is zero?

(c) Is there ever a dc current flowing through R_L if the input is sinusoidal?

(d) What would happen to the current flow through R_L over a complete cycle if the upper secondary winding were reversed in phase? Would there be any detrimental effect on the transistors?

Q20-8. Explain how the two complementary transistors in Fig. 20-11 are acting simultaneously as phase inverters and an output push-pull pair.

Q20-9. Refer to Fig. 20-13.

(a) What function do R_{12} and C_2 perform?
(b) Which of the transistors Q1 through Q4 provide current gain? Which provide voltage gain?
(c) What is the bias of Q2 determined by?
(d) What is the bias of Q1 determined by?
(e) What type of feedback is employed from the junction of R_{15} and R_{16} to Q1?
(f) What components determine the bias for Q3? How is feedback involved with R_9 and R_{10}?

Q20-10. (a) Explain what function the diode in Fig. 20-14 performs.
(b) How does it provide a temperature-compensating effect?
(c) How does the center-point voltage in Fig. 20-13 change when a positive going sine wave is applied to the input? Does the current through R_L increase or decrease?

PROBLEMS

P20-1. Consider the phase inverter in Fig. 20-5. $R_3 = R_4 = 1$ K ohm. The transistor's $h_{ie} = 2$ K ohms, and $h_{fe} = 100$.

(a) Determine the input resistance seen by v_s assuming $R_1 \parallel R_2$ is very large.
(b) Determine the two voltage gains, $\frac{v_{s1}}{v_s}$ and $\frac{v_{s2}}{v_s}$. Then obtain the output voltages for a 1 V p-p input.
(c) What is the approximate output resistance seen by v_{s1} and v_{s2}? ($R_s = 600$ ohms.)
(d) Draw the waveforms of v_{s1} and v_{s2} with $v_s = 1$ V p-p when the inverter is used to drive a class B push-pull amplifier, in which the load during conduction is assumed to be constant at 1 K ohm and during nonconduction is infinite. (Assume NPN transistors are used in the push-pull amplifier.)

P20-2. The conversion efficiency of a class B push-pull amplifier is being determined experimentally. An output power meter on the secondary of the output transformer indicates 5 W. The transformer is known to be 70% efficient. An oscilloscope connected across a 0.6-ohm resistor in series with the power supply shows a full-wave rectified waveform of peak value 380 mV. If $V_{CC} = 24$ V, determine the conversion efficiency of the amplifier.

P20-3. A class B push-pull amplifier must deliver 10 W of audio power to the output load.

(a) If the output transformer is 75% efficient, what is the minimum power drain on the power supply under optimum conditions?
(b) If the amplifier is modified to provide class AB operation at an efficiency of 68%, what will be the new power requirement?
(c) What is the minimum average dissipation capability required for each transistor for the conditions in (a)?
(d) If the same power were required from a single-transistor transformer-coupled amplifier as in (a), what would the transistor's average dissipation capability have to be?

P20-4. Design a class AB amplifier, as in Fig. 20-9, to provide 9 W of power in an 8-ohm load. Two 2N2869 germanium transistors are to be used, each having $P_T = 30$ W at flange temperatures up to 55°C. Since it is to be used in an automobile, use $V_{CC} = 14$ V and $R_e = 0.1$ ohm; h_{ie} for each transistor is 10 ohms, and $h_{fe} = 100$. Specify the output and input transformers (75% efficient) on the basis of their impedance matching capabilities and the values of R_1 and R_2. Assume $R_s = 600$ ohms.

P20-5. (a) Specify the input transformer for P20-4 with the emitter resistors not used.

(b) Determine the conversion efficiency. Can this level of efficiency be expected in practice? Why not?

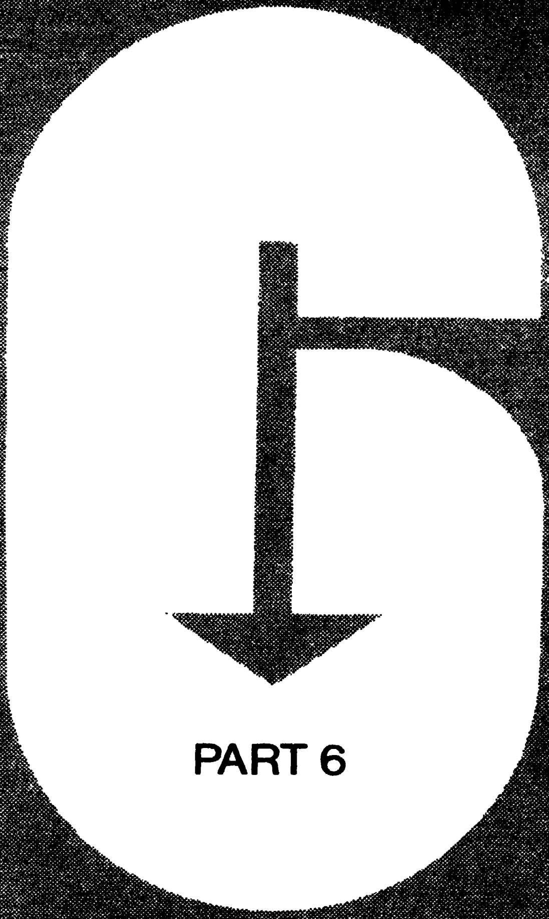
P20-6. Two 2N1701 silicon power transistors are to be used in a class AB push-pull audio amplifier, as in Fig. 20-9, to provide 15 W of power in a 4-ohm load. $h_{fe} = 50$, $P_T = 25$ W and $h_{ie} = 20$ ohms. No emitter resistors are used, and $V_{CC} = 24$ V. Specify the input and output transformers, the bias resistors, and the power drain on V_{CC}. Assume the output transformer is 75% efficient, and $R_s = 600$ ohms.

P20-7. Consider the amplifier in Fig. 20-10.

(a) What is the forward bias provided by resistors R_1-R_2 and R_3-R_4 with zero signal?
(b) Assuming the input signal drive to the base of Q1 and Q2 is such as to provide a minimum collector-to-emitter voltage of 7 V, what peak collector current occurs in each transistor?
(c) Determine the peak-to-peak load current and voltage, and hence the output power.
(d) What is the *total* dc power provided by both V_{CC} supplies?
(e) Determine the conversion efficiency.

P20-8. Refer to Fig. 20-13.

(a) If the center point voltage is 22 V, what are the relative resistance values of Q5 and Q6?
(b) What must the effective resistance of Q5 decrease to so as to cause the center-point (C.P.) to rise to 44 V?
(c) If the circuit is modified by the addition of a diode as in Fig. 20-14, what will be the base voltages when C.P. = 22 V?
(d) If an increase in base current for Q6 occurs (due to an input signal), what happens to the C.P. voltage?



PART 6

FIELD EFFECT TRANSISTORS

21

FIELD-EFFECT TRANSISTOR AMPLIFIERS

21-0. INTRODUCTION

The field-effect transistor (FET) may be thought of as the solid-state counterpart of a pentode; its characteristic curves are very similar. There are two types of FET, as shown in Figure 21-1. The first was described in 1952 by Shockley, but not made available until ten years later because of fabrication problems. It is the junction field-effect transistor, the JFET, which is a three-terminal device and is the one we shall concentrate on in this chapter. The second type is the insulated-gate field-effect transistor, IGFET, more commonly known as the metal-oxide-semiconductor (MOS) transistor or MOSFET. First described in a 1925 Canadian patent application, the IGFET was experimented with by Shockley *et al* before the bipolar junction transistor. The IGFET is a four-terminal device (although it is used basically as a three-terminal device in most applications) whose fabrication is very closely related to integrated circuit technology. This is discussed in detail in Chapter 22.

21-1. ADVANTAGES AND DISADVANTAGES OF THE FET

The main advantage of the FET is its high input resistance, in the order of 100 M ohms for a JFET and 10^{10} to 10^{15} ohms for the MOSFET. Thus, the FET is a voltage-controlled device, like a tube, and not current-controlled, like a conventional transistor. It shows a high degree of isolation between input and output.

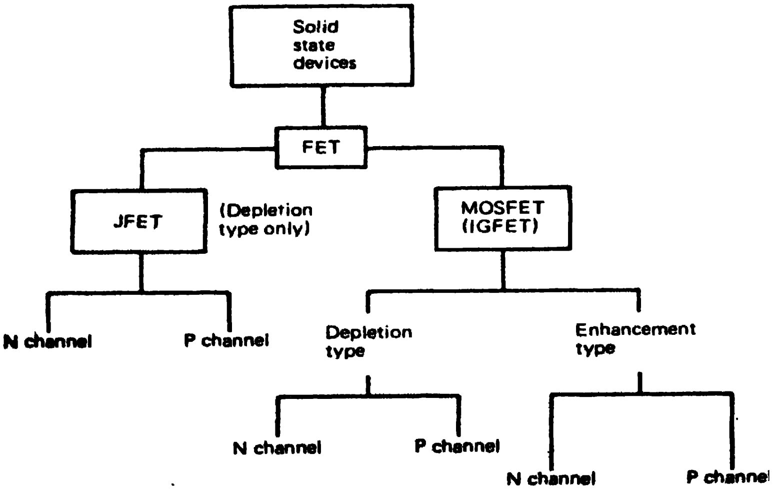


Fig. 21-1. FET "tree" showing the different types available.

The FET is a unipolar device, depending only upon majority current flow (either electrons in the N-channel type *or* holes in the P-channel type). The conventional transistor is a bipolar device relying upon two types of charge carrier, electrons *and* holes.

The FET is less noisy than a tube or bipolar transistor and is thus found in front ends of FM tuners for quiet reception.

The main disadvantage of the FET is the relatively low gain-bandwidth product compared with what can be achieved in conventional transistors. The MOSFET has the drawback of being very susceptible to overload voltages and may even require special handling during installation.

21-2. BASIC CONSTRUCTION OF THE JFET

A cross-sectional view of an N-channel JFET is shown in Fig. 21-2 together with the supply voltages and circuit symbol. (This is not an amplifier, but is meant to show the normal application of bias voltages and how the characteristic curves may be obtained.)

The symbol and structure bear a striking resemblance to the uni-junction transistor. The main difference is the FET is operated with the PN junction reverse biased, and therefore the gate draws negligible current. Note that the symbol has an arrow pointing in the direction in which the conventional current would flow if the junction were forward biased A P-channel JFET would have the arrow pointing out, similar to an NPN transistor.

Basically, the device consists of a lightly doped N-type *channel* diffused into a P-type bulk or substrate. (There is no electrical connection made to the bulk in a JFET). There is an ohmic connection (nonrectifying junction) made to one end of the channel called the *source*, through which the majority carriers (electrons in this case) enter the channel. They flow through the N-type bar and out through the *drain*. A heavily doped

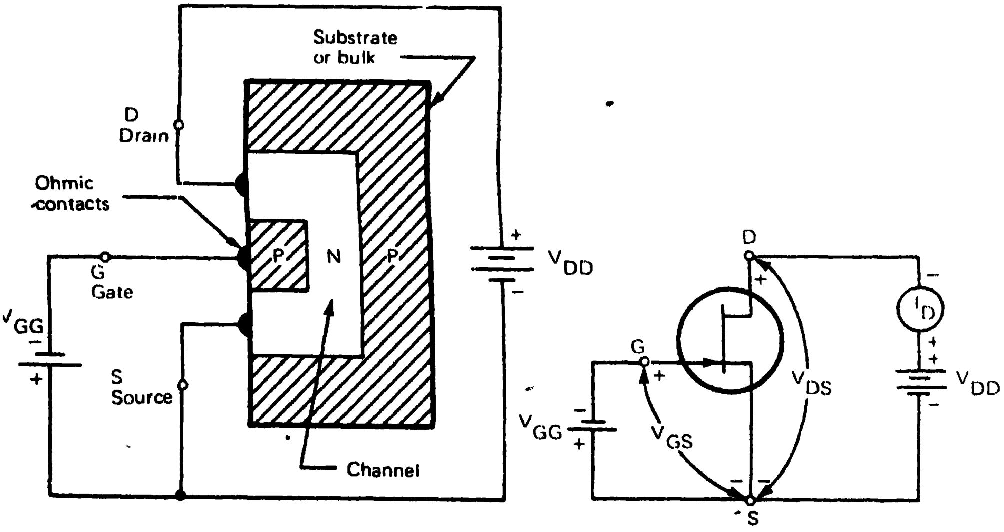


Fig. 21-2. (a) Diagram of a JFET with normal bias voltages. (b) Symbol of the N-channel JFET.

P-type concentration, the *gate* is diffused into the N channel, creating a PN junction. This is the terminal or electrode which will control the flow of drain current I_D, from S to D.

21-3. CHARACTERISTIC CURVES OF THE JFET

It can be seen from the preceding diagram that application of a voltage V_{DS} from drain to source will cause electrons to flow through the channel. The amount of current I_D will be determined, initially, by the value of V_{DS}, since it is just the ohmic resistance of the bar from S to D that is limiting current. But what effect will a voltage V_{GS} from gate to source have on this drain current for a given drain-to-source voltage?

Let us assume first that $V_{GS} = 0$ V. (The gate is shorted to the source.) The PN junction is not reverse biased. As V_{DS} is increased, drain current will increase also in a nearly linear manner until, at some voltage between 2 V and 8 V, a saturation effect is noted, and a "knee" develops in the characteristic curves as shown in Fig. 21-3.

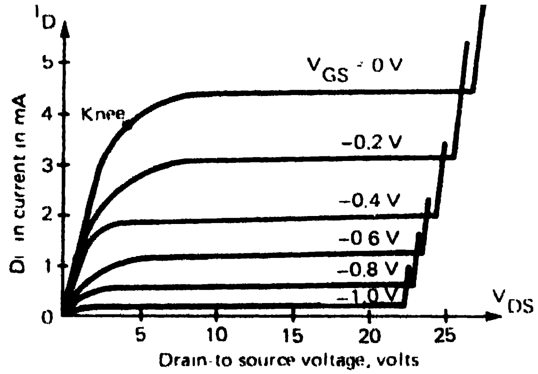


Fig. 21-3. Typical characteristic curves of an N-channel FET.

Little increase in drain current is observed with increasing drain-to-source voltage until an avalanche effect due to diode breakdown is caused (to be explained later).

If now the gate is made negative with respect to the source and the drain-source voltage is once again varied from zero up to breakdown, the curve is of a similar shape but shifted down. Apparently, a negative voltage on the gate has reduced drain current, much as a negative voltage on the grid of a tube will decrease plate current. However, the operation of the FET cannot be explained as the tube is explained, by just considering electrons, and saying that the negative gate *repelled* electrons, since it is possible to have a P-channel FET, in which a positive gate voltage will also reduce drain current. See Fig. 21-4.

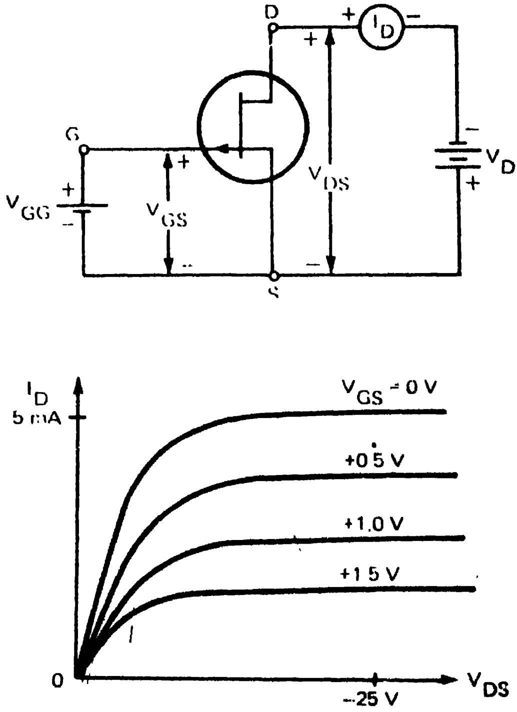


Fig. 21-4. Symbol and characteristic curves for a P-channel FET.

However, in the P-channel type, in which holes (positive) are the current carriers, a positive gate could be considered as repelling holes.

21-4. PRINCIPLES OF OPERATION OF THE JFET

Actually, the control of drain current through the channel is not a repelling effect but a constricting or narrowing of the channel effect. How does the gate achieve this?

Consider what happens to the PN junction between gate and source under reverse biased conditions, as depicted in Fig. 21-5.

Assume for the present no voltage is applied between D and S. With a small reverse bias, say $V_{GS} = 1$ V, the depletion region at the junction is quite narrow. Electrons have been drawn from the junction in the N-type material, leaving bound, immobile positive ions in that region of the channel, effectively reducing the width and conductivity of the channel. (Similarly negative ions are left in the P-type material near the junction.) As the reverse bias is increased to -3 V and -6 V, the depletion region or space-charge region widens, reaching ever further into the channel and restricting the passage of electrons from S to D when a voltage is applied. Finally, a value of voltage from gate to source can be applied that *pinches off* channel conduction altogether and cuts off drain current. This value of gate voltage is appropriately called the *pinch-off* voltage, V_P, and is referred to in the specifications as $V_{GS(off)}$. For example, a 2N3819 N-channel FET has a $V_{GS(off)}$ of -8 V, at which $I_D = 2$ nA (nanoampere) when $V_{DS} = 15$ V.

The mechanism of current control is clearly dependent upon the degree to which the electric *field* (due to the charge in the depletion region) extends into the channel and provides the *effect* of decreasing the conductivity through the *transistor*. Hence the name of *field effect transistor*, or FET.

21-5. EFFECT OF V_{DS} ON CHANNEL CONDUCTIVITY

The depletion region is not symmetrical around the PN junction when a voltage is applied from drain to source, but is wider (W is smaller in the previous diagrams) the closer one moves toward the drain. This is explained by considering the answer to the following question.

Why, when V_{GS} was zero (no reverse bias on the gate), did the drain current I_D begin to level off in the characteristic curves when V_{DS} exceeded a certain value?

Consider the JFET with G shorted to S and a very restricted model of the device, as shown in Fig. 21-6.

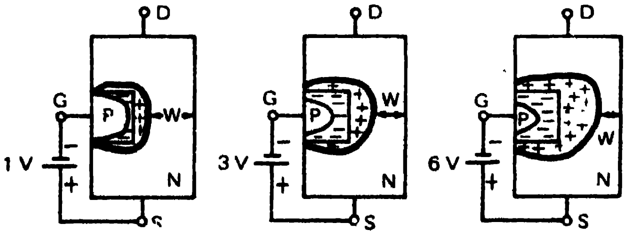


Fig. 21-5. Effect of reverse bias on the width of the channel

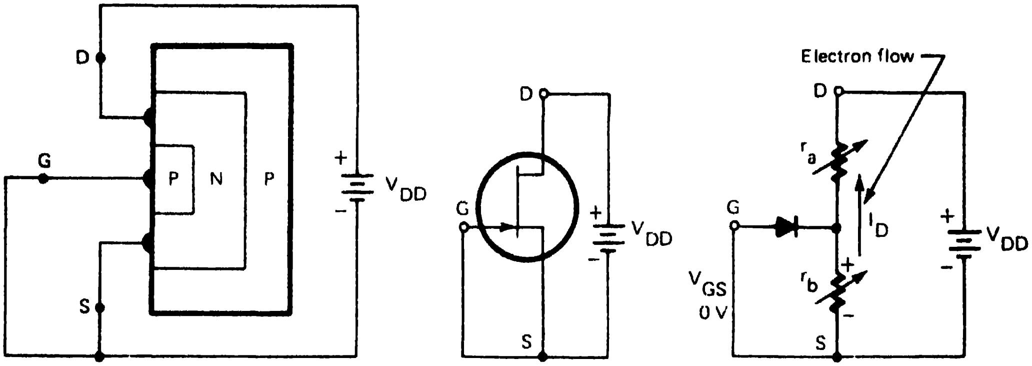


Fig. 21-6. N-channel FET with a simplified model.

Resistances r_a and r_b represent the effective resistance of the channel. They are shown variable since their values depend upon both V_{DS} and V_{GS}.

Referring to the model of the JFET (valid only for this discussion), it can be seen that the IR voltage drop caused by the drain current through the channel resistance has the effect of reverse biasing the gate-source PN junction and causing a constriction just as explained earlier. The pinching effect increases with increase in I_D (and V_{DS}) until a condition of *self pinch-off* occurs, where there is no further appreciable increase in drain current. (The channel does not cut off completely, since this would reduce I_D to zero and remove the reverse-biasing effect).

Due to the progressive voltage drop along the length of the channel, this reverse biasing effect is strongest near the drain end of the gate, so that a wedge-shaped depletion region is created. See Fig. 21-7.

If the drain-source voltage exceeds the reverse breakdown voltage of the PN junction from drain to gate (25 V in the 2N3819), a large current I_D will flow, which must be avoided in actual operation. It is to be noted from the characteristic curves that as V_{GS} is applied (to reverse bias the gate) breakdown due to V_{DS} gets successively smaller by the amount of V_{GS}. This is because reverse biasing of the PN junction is due to the combined efforts of both the externally applied voltage V_{GS} and the inherent self pinching-off action due to the IR drop within the channel.

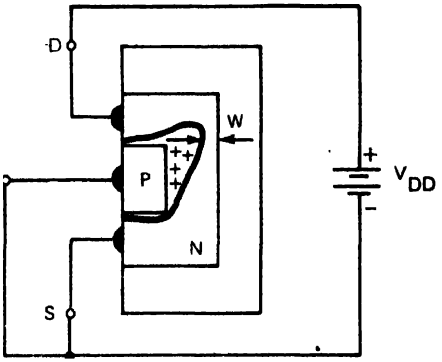


Fig. 21-7. Showing the unsymmetrical depletion region in the channel due to I_D

21-6. CHANNEL OHMIC REGION AND PINCH-OFF REGION

If on the characteristic curves a vertical line is drawn where V_{DS} is numerically equal to V_P, the pinch off voltage [$V_{GS(off)}$], two regions are defined as shown in Fig. 21-8.

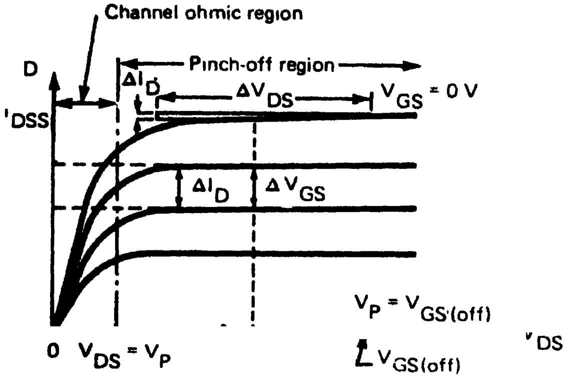


Fig. 21-8. Typical characteristic curves showing the channel ohmic and pinch-off regions.

In the channel ohmic region for small values of V_{DS}, the resistance from drain to source, $R_{DS} = \frac{V_{DS}}{I_D}$, is related to the gate voltage, V_{GS}, in an almost linear manner. This makes it most useful as a voltage variable resistor (VVR) or voltage dependent resistor (VDR), which will be considered later in Chapter 22. The pinch-off region is reserved for high-impedance voltage amplifiers and can be used for a graphical analysis of such amplifiers.

Before proceeding to the use of the FET as an ac amplifier, let us look at the parameters of the FET, obtainable from the characteristic curves or given in the specifications. (These parameters are analogous to the μ, r_p, and g_m of the pentode, as covered in Appendix D. If they are found to have a similar relationship, then all the equations developed for the triode and pentode concerning impedances and voltage gains will equally apply, and only a summary of the different configurations will be necessary.)

21-7. CHARACTERISTIC PARAMETERS OF THE FET

Drain Resistance, r_d

Corresponding to the plate resistance of a tube, r_p, we can determine the drain resistance, r_d, defined as follows:

$$r_d = \frac{\Delta V_{DS}}{\Delta I_D}, \; V_{GS} \text{ held constant.} \qquad (21\text{-}1)$$

(See previous characteristic curves in Fig. 21-8.) This is usually evaluated (as far as the specification sheet is concerned) at $V_{GS} = 0$ V and come prescribed drain voltage around 10 to 15 V. Because of the flatness of the characteristic curves, this may be difficult to do graphically, but can be readily measured experimentally.

The specification sheet may give the drain (or output) resistance r_d directly in ohms, but some manufacturers provide the parameter y_{os}, which is an admittance quantity (mhos) and is a measure of both the output resistance r_d and output capacitance. At low frequencies then, when the capacitive effect is negligible

$$1 \qquad (21\text{-}2)$$

For example, the 2N3819 has a typical maximum value of y_{os} at 1 kHz of 50 μ mhos. Thus,

$$r_d = \frac{1}{50 \times 10^{-6}} = 20{,}000 \text{ ohms.}$$

This is typical for many JFETs, although the output resistance, sometimes referred to as r_{oss}, varies widely with drain current and from one device to another of the same type.

Transconductance, g_{fs}

The control that the gate voltage has over the drain current is measured by the forward transconductance, g_{fs}, and is similar to the mutual conductance, g_m, of a tube

$$g_{fs} = \frac{\Delta I_D}{\Delta V_{GS}}, \; V_{DS} \text{ held constant.} \qquad (21\text{-}3)$$

This is also usually evaluated in the region where $V_{GS} = 0$ and $V_{DS} = 15$ V, (but in fact g_{fs} decreases almost linearly with an increase in V_{GS}). The specification sheet will provide the forward transfer admittance, y_{fs}, which also includes the capacitive effect of the device. But at the low frequencies this effect is negligible, so

$$g_{fs} \simeq y_{fs} \qquad (21\text{-}4)$$

For the 2N3819, y_{fs} has a minimum value of 2000 and a maximum value of 6500 μ mhos. Here again we see the typically large variation that occurs in solid-state devices.

Amplification Factor, μ

It can be shown that the amplification factor, μ, of an FET is related to r_d and g_{fs} by

$$\mu = r_d g_{fs} = \frac{\Delta V_{DS}}{\Delta V_{GS}}, \; I_D \text{ held constant,} \qquad (21\text{-}5)$$

which for the 2N3819 is between 40 and 130. This is intended to show how much more control the gate voltage has over drain current than has the drain voltage.

Zero-Gate-Voltage Drain Current, I_{DSS}

To give some idea of the drain current capability, the specification sheet gives I_{DSS}, the zero-bias drain current (current from drain to source with the gate shorted). For the 2N3819, I_D is evaluated at a V_{DS} of 15 V and may be anywhere between 2 and 20 mA. This is an important quantity, since if the voltage used to reverse-bias the gate were ever lost (V_{GS} = 0 V), then the drain current would jump to the value of I_{DSS} if the drain-to-source voltage exceeded V_P. This may cause overheating of the device.

Dissipation, P_D

The device's continuous dissipation, P_D, is the product of I_D and V_{DS}, and for the 2N3819 has a maximum value of 200 mW. This is quite typical, but must be reduced for temperatures above 25° C.

Gate Cutoff Current, I_{GSS}

It has been shown that the device is controlled by a reverse-biased PN junction. Thus, the gate current, which determines the input resistance of the FET, is understandably small. The gate cutoff current, I_{GSS}, for a JFET is typically a few nanoamperes. (1nA = 1×10^{-9} A).

For the 2N3819, I_{GSS} is 2 nA when V_{GS} = −15 V and V_{DS} = 0. This represents an input resistance from gate to source (sometimes called gate-to-source leakage resistance) of

$$r_{gs} = \frac{V_{GS}}{I_{GSS}} = \frac{15}{2 \times 10^{-9}} = 7.5 \times 10^9 \text{ ohms}$$
$$= 7{,}500 \text{ M ohms.} \qquad (21\text{-}6)$$

21-8. EFFECT OF TEMPERATURE ON FET PARAMETERS

All the parameters above were specified at a temperature of 25° C; but, being a silicon device, the FET can withstand operation up to 125° C, provided that the changes that take place in the parameters are taken into account.

As the temperature increases, g_{fs} decreases (it is 40% lower at 150° C), as does the drain current I_D. As can be expected, the drain resistance increases. What is the reason for the negative temperature coefficient for I_D? Recall that drain current is due only to majority carriers, so their mobility is decreased at elevated temperatures due to the greater collisions that take place between them and the remaining ions in the semiconductor channel. (Much the same reason as the increase in resistance of copper wire with increased temperature.) Thus, thermal runaway is no problem in an FET. However, where the drain current is used to self bias an FET, this variation of I_D with temperature *is* troublesome; it is considered in Section 21-10.

Actually, there is a second factor that affects the variation of drain current with temperature. An increase in temperature decreases the width of the channel constriction, which allows I_D to increase. This effect becomes dominant at low values of drain current (high gate-to-source reverse voltages) so that an *increase* in drain current results from an increase in temperature. See Fig. 21-9(a). These transfer curves show how the drain current varies with temperature and gate-to-source voltage. They also indicate the minimum and maximum values of drain current that can occur for this transistor (I_{DSS} has a minimum of 7 mA and a maximum of 14 mA at 25° C).

Note that for gate-to-source voltages in excess of 5.5 V the drain current shows a slight increase with temperature instead of a decrease. It is thus possible to bias the field-effect transistor at a gate-to-source voltage that would provide zero drain-current drift with temperature variation. However, the transconductance is much reduced in this region, as can be seen in Fig. 21-9(b). The transconductance is approximately 1500 μmhos at I_D = 0.5 mA, compared with 3000 μmhos at 4 mA, with the result that voltage gain suffers considerably.

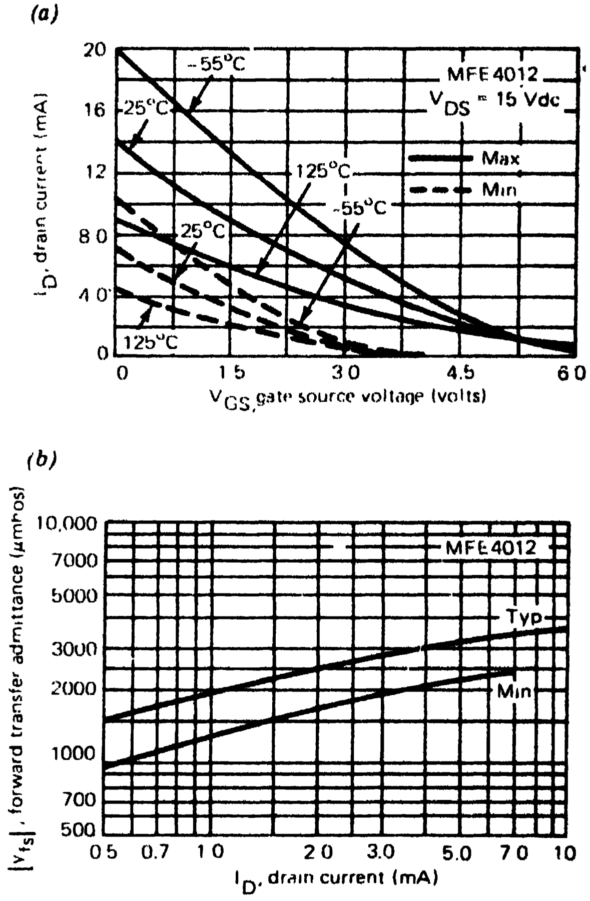


Fig. 21-9. (a) Transfer curves of a P-channel, type 4012 junction field-effect transistor showing how the drain current varies with gate-to-source voltage and temperature. (b) Showing how y_{fs} varies with drain current. (Courtesy of Motorola Inc.)

The gate current does not have a negative temperature coefficient and shows a large variation with temperature. For the 2N3819, I_{GSS} is 2 nA when V_{GS} is -15 V and the temperature is 25° C, but increases to 2 μA at 100° C, - a factor of 1000. Of course, the reason for this is the reverse biasing of the PN junction, which means that the current is composed of minority carriers, which increase in mobility with temperature as explained in Chapter 5.

21-9. COMMON-SOURCE AC AMPLIFIER

The similarity of the FET to the tube has been indicated. Analysis and design may be carried out graphically or by using the same equivalent model as a triode (or pentode) to develop equations for voltage gain, output impedance, etc. The only difference is the large variation in characteristic curves from one unit to the next and their variation with temperature.

Shown in Fig. 21-10 is a common-source ac amplifier, with typical characteristic curves of the FET.

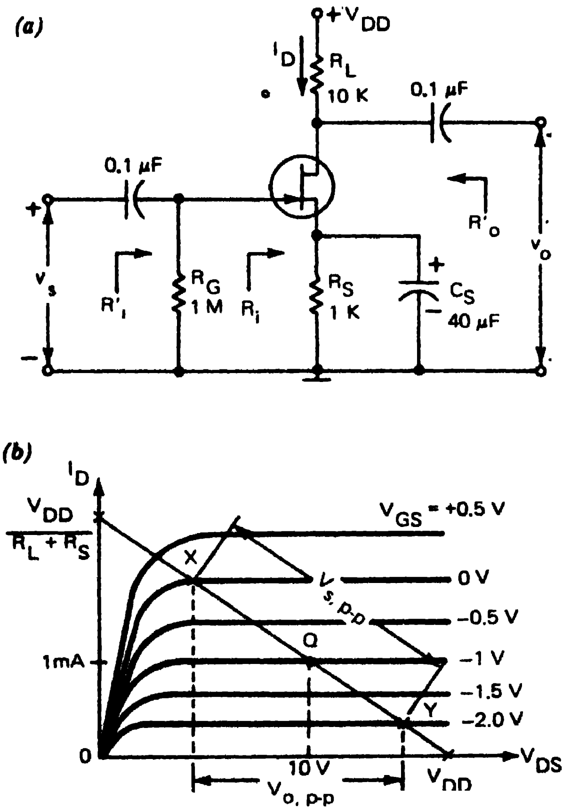

(a)
+V_DD
I_D
R_L
10 K
0.1 μF
0.1 μF
R'_o
v_o
+
v_s
−
R'_i
R_G
1 M
R_i
R_S
1 K
+
C_S
40 μF
(b)
I_D
V_DD / R_L + R_S
V_GS = +0.5 V
X
0 V
−0.5 V
Q
1 mA
−1 V
−1.5 V
Y
−2.0 V
0
10 V
V_DD
V_DS
V_o, p-p

Fig. 21-10. (a) Common-source FET ac amplifier. (b) Characteristic curves with dc load line and Q point.

Note that I_D (arrow shown for conventional current flow through FET) flows through R_S to develop a voltage that reverse-biases the PN gate junction. R_G is provided for a dc return path.

The drawing of the dc and ac load lines proceeds as for a transistor amplifier. The Q point is found by trial and error. It is the point on the ac load line where the bias equation $V_{GS} = -I_D R_S$ is satisfied. The voltage gain may be found from the ac load line and is given by

$$\frac{v_{ds}}{v_{gs}} = \frac{v_o}{v_s} \tag{21-7}$$

It is typically 10 to 50.

Voltage Gain, A_v

Since the same equivalent circuit can be used as for the triode, the voltage gain equation will be the same and is repeated here using FET terminology. Refer to Fig. 21-11.

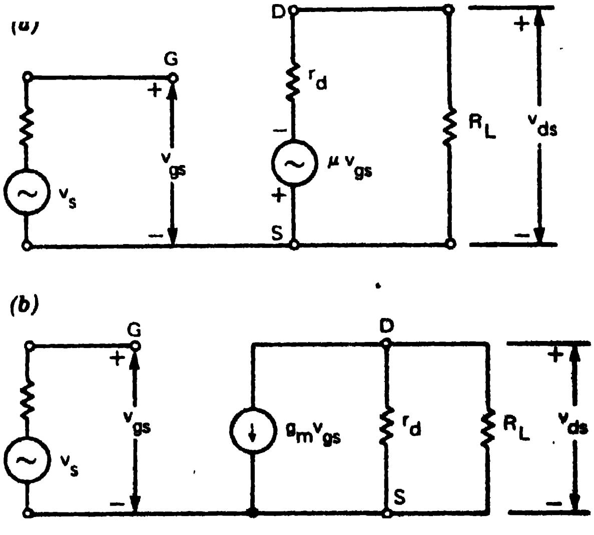

(a)
G
+
v_gs
−
v_s
D
r_d
−
μ v_gs
+
S
R_L
+
v_ds
−
(b)
G
+
v_gs
−
v_s
D
g_m v_gs
r_d
S
R_L
+
v_ds
−

Fig. 21-11. (a) Thevenin equivalent circuit for small-signal common-source ac amplifier. (b) Norton's current equivalent circuit.

$$A_v = \frac{-\mu R_L}{r_d + R_L} = -g_{fs}\frac{r_d R_L}{r_d + R_L} = -g_{fs}\frac{R_r}{1 + \frac{R_r}{\cdot}} \qquad (21\text{-}8)$$

If $r_d >> R_L$ (not always the case), then the same equation holds true as for the pentode

$$A_v \simeq -g_{fs}R_L. \qquad (21\text{-}9)$$

Using the minimum and maximum values of g_{fs} for the 2N3819 (2000 to 6500 μmhos), R_L = 10 K ohms, and r_d = 20 K ohms, gives an A_v of -13 to -42.

Output Resistance, R_o'

From the equivalent circuit, $R_o' = R_L \| r_d = \frac{R_L r_d}{R_L + r_d}$, (21-10)

which is approximately 7 K ohms in the above example.

Input Resistance, R_i'

It has been shown that the input resistance between gate and source, r_{gs}, is typically in excess of 100 M ohms. Thus, the input resistance to the amplifier is determined by R_G, and in this case is 1 M ohm. Thus, for a high input impedance amplifier, R_G should be as large as possible, but there are two restraints on this.

(1) The noise figure of the amplifier will be at a minimum when R_G is in the order of slightly less than 1 M ohm (noise figure of the FET is around 2 to 4 dB).

(2) The reverse leakage current that flows through the gate-to-source PN junction, although small, may develop sufficient voltage to shift the operating point significantly, especially at higher temperatures. This can be seen by referring to the restricted equivalent circuit used earlier and repeated in Fig. 21-12.

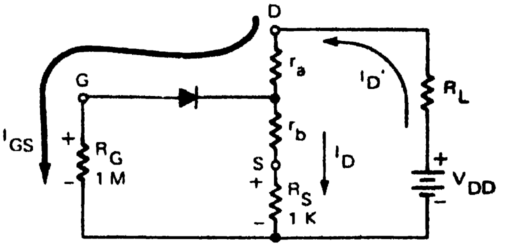


Fig. 21-12. Simple equivalent circuit showing effect of leakage current I_{GS}.

Whereas the drain current serves to reverse bias the junction, the gate current, I_{GS}, develops a voltage across R_G that forward biases the junction. If I_{GS} is 2 nA (at 25°C) and R_G is 1 M ohm, then the voltage developed accross R_G is only 0.002 volt. But, if I_{GS} increases to 2 μA at 100°C, then the voltage developed will be 2 volts and will reduce the negative bias to

$$V_{GS} = -I_D R_S + I_{GS}R_G. \qquad (21\text{-}11)$$

$$V_{GS} = -1 + 2 = +1 \text{ volt for } I_D = 1 \text{ mA}.$$

The FET will be forward biased, which may cause distortion and may also cause overheating of the gate. Special compensation could be used where high temperatures would be encountered, but the above calculation does indicate that it is not prudent to increase R_G in an attempt to increase input resistance. If a higher input impedance is required, a common-drain or source follower would be used. (See Section 21-11.)

Input Capacitance, C_i

The FET also has "interelectrode" capacitances (due to the PN junctions and channel) which limit the frequency response of an amplifier connected in the common source mode. Fig. 21-13.

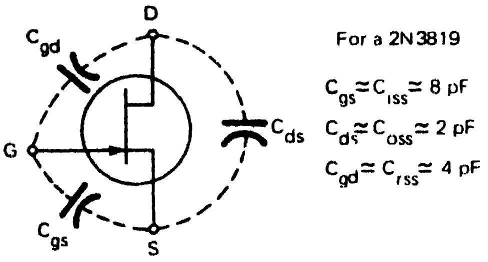

D
G
S
C_{gd}
C_{ds}
C_{gs}
For a 2N3819
$C_{gs} \simeq C_{iss} \simeq 8$ pF
$C_{ds} \simeq C_{oss} \simeq 2$ pF
$C_{gd} \simeq C_{rss} \simeq 4$ pF

Fig. 21-13. Diagram showing the capacitances in an FET with typical values for a 2N3819.

The triple subscript parameters are those given in the specification sheet and represent the common-source short-circuit capacitances.

The Miller effect increases the input capacitance over the value C_{gs}, as given by

$$C_i = C_{gs} + (1 - A_v)C_{gd},$$

$$\text{where } A_v = -g_{fs} \times R_L \parallel r_d. \qquad (21\text{-}12)$$

It is the large value of C_{gd} (C_{rss}) that is harmful in high frequency work, (especially in the operation of cascaded amplifiers as in Section 21-13), but the MOSFET typically has a lower value. An N-channel 40467 MOSFET has 0.05 pF for C_{rss} and is capable of working in VHF amplifiers around 200 MHz when neutralized (to reduce the feedback capacitance effect).

21-10. FIXED BIAS WITH SELF-BIAS

The common source ac amplifier studied thus far only had self bias. *Question*: What would be the effect of replacing an FET which had a low value of drain current, say 2 mA, at zero bias, with an FET that had a large value of drain current, say 20 mA, at zero bias? (Zero bias is used just for some reference.)

Answer: If the source resistance (the resistor in the source lead of the FET, which is usually quite small, say 1 K ohm) is unchanged, then the increase in drain current will cause only a small increase in negative bias, not sufficient to bring the drain current and operating point back to its initial value. If in some way the source resistance could be made larger (even with the low drain current transistor in place), then a change in drain current (upon changing transistors or temperature) would cause an appreciable change in bias, which would tend to restore the status quo (initial operating point).

This is achieved in the circuit shown in Fig. 21-14.

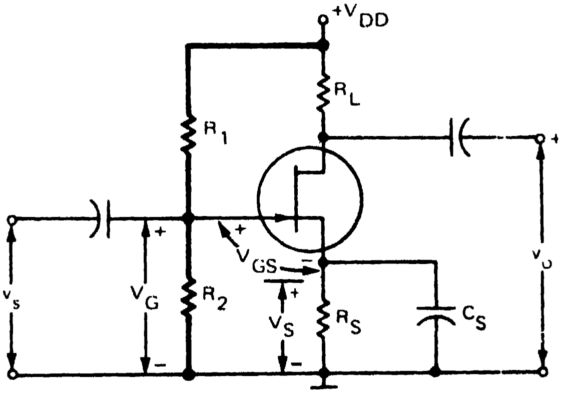


Fig. 21-14. Common-source ac amplifier with additional bias due to R_1 and R_2.

Note that the voltage divider $R_1 R_2$ applies a fixed forward bias to the gate and that this must be offset by enlarging R_S and its resulting negative self-bias. This circuit also helps to bias against drift of drain current under temperature variations, as shown by the following example.

Example 21-1. Let it be required to select bias resistors R_1, R_2, and R_S in Fig. 21-14 to provide a quiescent drain current of 4 mA. Assume that the transfer characteristic curves of Fig. 21-9 apply to the transistor and that the drain current must remain within ±25% of 4 mA over a temperature range from -55°C to 125°C. Take into account the maximum and minimum (worst-case) drain currents that the transistor may have. $V_{DD} = 20$ V.

Solution. The transfer curves have been reproduced in Fig. 21-15 to show the extreme limits arising from device and temperature variation. The maximum and minimum permissible drain currents are

$$I_{D,max} = 4 + 1 = 5 \text{ mA}; \quad I_{D,min} = 4 - 1 = 3 \text{ mA}.$$

Plot these values on the curves and establish the maximum and minimum gate-to-source voltages

$$V_{GS,max} = -3.62 \text{ V} \qquad V_{GS,min} = -0.78 \text{ V}$$

Now refer to Fig. 21-14 and note that the gate-to-source bias voltage is given by

$$V_{GS} = V_G - V_S = V_G - I_D R_S. \tag{21-13}$$

This equation is represented by the bias line drawn in Fig. 21-15. That this is so can be seen by rearranging Eq. 21-13

$$I_D = \frac{-1}{R_S} V_{GS} + \frac{V_G}{R_S}. \tag{21-14}$$

Thus, the bias line has a slope equal to $\frac{1}{R_S}$, which permits us to find the value of R_S from the following

$$R_S = \frac{V_{GS,max} - V_{GS,min}}{I_{D,max} - I_{D,min}} \tag{21-15}$$

which yields 1.42 K ohms for our example.

Examination of Eq. 21-14 shows that I_D will be zero when $V_{GS} = V_G$. Hence, the intercept of the bias line on the V_{GS} axis provides $V_G = +3.5$ V.

Alternatively, the value of V_G may be obtained from the bias line equation

$$V_G = \frac{I_{D,max} V_{GS,min} - I_{D,min} V_{GS,max}}{I_{D,max} - I_{D,min}} \tag{21-16}$$

$$= +3.48 \text{ V}.$$

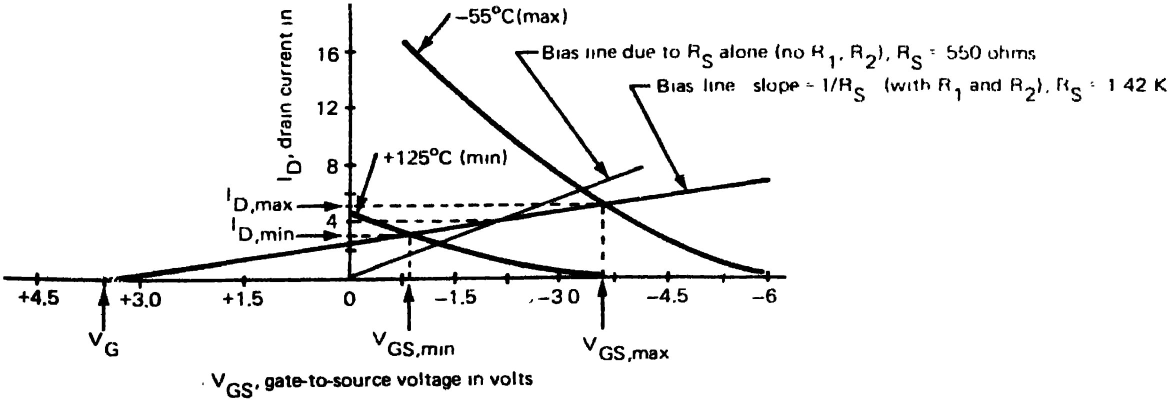


Fig. 21-15. Extreme transfer curves for -55°C and +125°C with the bias line drawn between the maximum permissible drain current limits.

Now V_G must be developed across R_2 by R_1, R_2 and V_{DD}. Thus, ignoring the gate reverse current

$$R_1 + R_2 \quad {}'_{DD}\cdot \qquad (21\text{-}17)$$

Using $V_G = 3.48$ V and $V_{DD} = 20$ V, $R_1 = 4.75\ R_2$. The choice of R_1 and R_2 is limited by the gate reverse current, as covered in Section 21-9. It is generally sufficient to restrict the size of R_1 and R_2 so that approximately five times as much bias current flows as reverse gate current. Assuming the reverse gate current is $2\ \mu A$ at 125°C, we could make $R_1 = 1.6$ M ohms and $R_2 = 330$ K ohms.

Note that the gate-to-source bias voltage to make $I_D = 4$ mA is

$$v_{GS} \quad V_G - I_D R_S$$

$$= 3.48 - 4 \times 10^{-3} \times 1.42 \times 10^3$$

$$= 3.48 - 5.68 = -2.2 \text{ V}.$$

This is confirmed by the intersection of $I_D = 4$ mA and the bias line drawn in Fig. 21-15.

If a single resistor, R_S, were used to provide this bias (no R_1 and R_2), the value of R_S would be

$$\frac{2.2}{4 \times 10^{-3}} = 550 \text{ ohms}.$$

The bias line representing this condition would be a line drawn through the origin (Since $V_G = 0$) and the desired Q point at $I_D = 4$ mA and $V_{GS} = -2.2$ V. (See Fig. 21-15.) Note that the line is steeper (smaller R_S), causing more extreme values for $I_{D,max}$ (6.2 mA) and $I_{D,min}$ (2.45 mA). These values would be even more extreme if the required quiescent drain current were higher than 4 mA.

The effect of R_1 and R_2 is to stabilize the bias point against device and temperature variation and thus provide a voltage gain that is more nearly constant.

21-11. THE COMMON-DRAIN OR SOURCE FOLLOWER

The basic circuit of a source follower is shown in Fig. 21-16.

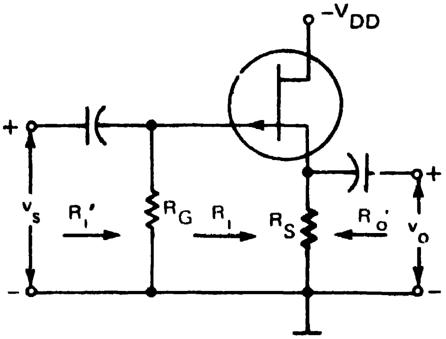


Fig. 21-16. Basic circuit of a common-drain amplifier or source follower.

Note that a P-channel FET has been used for circuit variation. Like the emitter follower, this circuit has a voltage gain of less than unity, no phase reversal, high input impedance and low output impedance.

Voltage Gain, A_v

$$A_v = \frac{R_S}{\underline{\mu + 1}\,R_S + \frac{1}{g_{fs}}} \quad \frac{g_{fs}R_S}{1 + g_{fs}R_S} \qquad (21\text{-}18)$$

The approximation holds when $\mu >> 1$, which is usually the case for an FET.

Output Resistance, R_o

$$R_o' = R_S || \frac{r_d}{1 + \mu} \simeq \frac{R_S}{1 + g_{fs}R_S} \text{ when } \mu >> 1. \qquad (21\text{-}19)$$

Input Resistance, R_i'

Although R_i (in Fig. 21-16) is extremely large, the overall amplifier input resistance, R_i', will be determined by R_G and will be typically 1 M ohm.

To increase the input resistance, the bootstrapping principle discussed in detail in Chapter 17 must be applied. A circuit to do this is shown in Fig. 21-17.

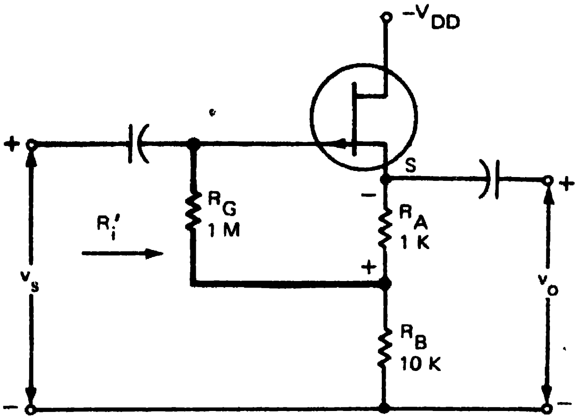


Fig. 21-17. Common-drain amplifier with bootstrapped bias circuit. ($R_S = R_A + R_B$)

Note that the fraction of the output voltage v_o fed back to R_G is

$$v_o \times \frac{}{R_A + R_B} \quad v_o \times \frac{10\text{ K}}{1\text{ K} + 10\text{ K}} \simeq 0.9\, v_o \tag{21-20}$$

so the bootstrapping will not be as effective as if R_G had been connected to the source directly. However, direct connection to the source would not be practical, since R_A is required to develop the necessary bias for the gate. The value of R_A is thus determined by the desired operating point, that is, I_D and V_{GS}, while the value selected for R_B will fect mostly the output resistance.

It may be shown that the input resistance for the above circuit is given by

$$R_i' \quad \frac{R_G}{1 - \dfrac{R_B}{R_A + R_B} A_v} \quad \frac{R_G}{1 - \dfrac{R_B}{R_S} A_v} \tag{21-21}$$

$$\text{If } R_B >> R_A,\text{ then } R_i' \quad \frac{R_G}{1 - A_v} \tag{21-22}$$

which is the same relationship as was developed for the bootstrapped Darlington circuit in Section 17-6.

In this example, assume $A_v = 0.98$ ($g_{fs} = 5000\ \mu$mhos).

$$\text{Then } R_i' \quad \frac{1\text{ M}}{1 - \dfrac{10\text{ K}}{1\text{ K} + 10\text{ K}} \times 0.98}$$

$$\frac{1\text{ M}}{1 - 0.88} = \frac{1\text{ M}}{0.12} \quad 8\text{ M ohms.}$$

If the approximation had been used

$$R_i' \quad \frac{1\text{ M}}{1 - 0.98} = \frac{1\text{ M}}{0.02} = 50\text{ M ohms.}$$

Thus, to increase the input resistance, R_B should be much larger than R_A, and A_v should be as close to unity as possible. However, there is a limit to the size of R_B, since a larger value of V_{DD} would be required for a given operating point. Also, R_o' would increase with increasing value of R_B (and R_S).

In this example, the output resistance would be, assuming $g_{fs} = 5000\ \mu$mhos,

$$R_o' \simeq \frac{R_S}{1 + g_{fs} R_S}$$

$$\frac{11 \times 10^3}{1 + 5000 \times 10^{-6} \times 11 \times 10^3} \approx 200\text{ ohms.}$$

The source follower is thus a good impedance-matching device.

Input Capacitance, C_i

$$C_i = C_{gd} + (1 - A_v)C_{gs}. \tag{21-23}$$

Thus, the input capacitance is reduced by the negative voltage feedback, since $A_v \simeq 1$.

Output Capacitance, C_o

$$C_o = C_{ds} + C_{gs}\,\frac{1 - A_v}{A_v}. \tag{21-24}$$

Similarly, the output capacitance is reduced to approximately C_{ds} when $A_v \simeq 1$.

21-12. THE COMMON-GATE FET AMPLIFIER

The circuit in Fig. 21-18 is the counterpart of the grounded-grid amplifier. Note that a dc path must exist through or around the signal voltage source, v_s and R_s, for biasing purposes.

Voltage Gain, A_v

$$\frac{V_o}{V_i} = A_v = \frac{(\mu + 1)R_L}{r_d + R_L} = \frac{(r_d g_{fs} + 1)R_L}{r_d + R_L} \tag{21-25}$$

Input Resistance, R_i

$$R_i = \frac{r_d + R_L}{\mu + 1} = \frac{r_d + R_L}{r_d g_{fs} + 1}. \tag{21-26}$$

Output Resistance, R_o

$$R_o = r_d + R_s(\mu + 1) = r_d + R_s(r_d g_{fs} + 1). \tag{21-27}$$

This amplifier provides a low input impedance, high output impedance, high voltage gain (if R_s is small), and no phase reversal. It is a desirable connection for high frequency work from a capacitive standpoint.

21-13. FREQUENCY RESPONSE OF THE FET AMPLIFIER

The frequency response of a cascaded common-source ac amplifier is similar to that employing pentodes, and the same equations apply as those developed in Appendix C. They are repeated below using FET terminology. Refer to Fig. 21-19.

Mid-Frequency Gain

$$A_{v(MF)} = \frac{V_o}{V_i} = -g_{fs}(r_d \| R_L). \tag{21-28}$$

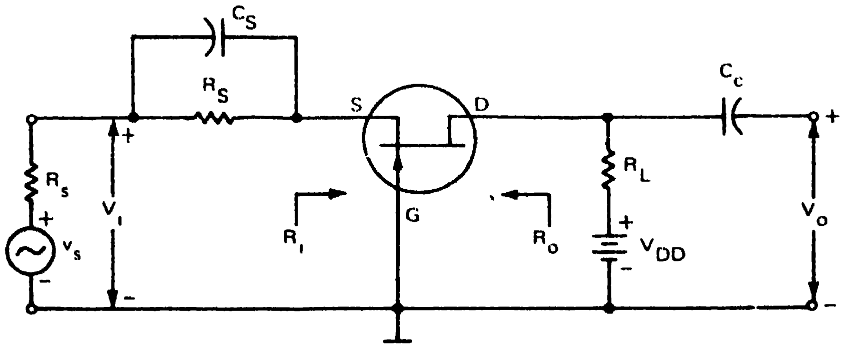


Fig. 21-18. Basic circuit of a common-gate amplifier.

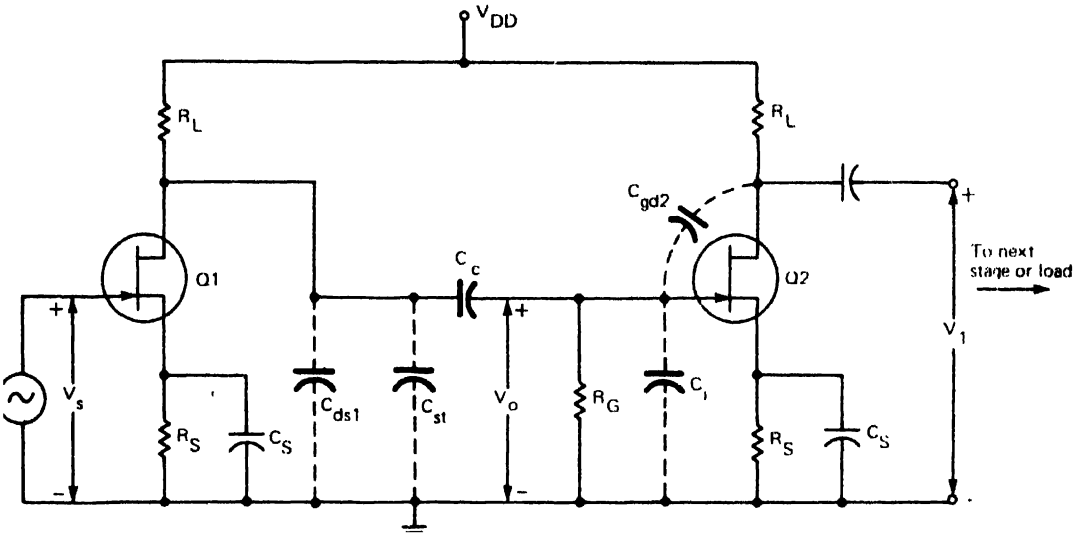


Fig. 21-19. Cascaded FET amplifier showing high-frequency capacitances.

High-Frequency Response

$$A_{v(HF)} = \frac{V_o}{V_s} = \frac{A_{v(MF)}}{1 + j\frac{R_{eq}}{X_{CT}}} \qquad (21\text{-}29)$$

$$R_{eq} = r_d \,||\, R_L \,||\, R_G \approx r_d \,||\, R_L. \qquad (21\text{-}30)$$

$$X_{CT} = \frac{1}{2\pi f\, C_T},\ C_T = C_{dsl} + C_{stray} + C_i. \qquad (21\text{-}31)$$

$$C_i = C_{gs2} + (1 - A_v)C_{gd2}, A_v = A_{v(MF)} = \frac{V_1}{V_o}. \qquad (21\text{-}32)$$

$$f_2 = \frac{1}{2\pi R_{eq} C_T}: \qquad (21\text{-}33)$$

Low-Frequency Response

(a) Assuming C_S is infinitely large (response dependent upon C_c),

$$A_{v(LF)} = \frac{V_o}{V_s} = \frac{A_{v(MF)}}{1 - j\frac{X_{Cc}}{R}}, \qquad (21\text{-}34)$$

where

$$X_{Cc} = \frac{1}{2\pi f C_c}, \text{ and } R = R_G + r_d \,||\, R_L \approx R_G. \qquad (21\text{-}35)$$

$$f_1 = \frac{1}{2\pi R C_c} \qquad (21\text{-}36)$$

(b) Assuming C_c is large (response depends upon C_S),

$$A_{v(LF)} = \frac{A_{v(MF)}}{1 + g_{fs}R_S} \times \frac{1 + j\,f}{1 + j\frac{f}{f}}, \qquad (21\text{-}37)$$

$r_d >> R_L + R_S;$

where $f_o = \dfrac{1}{2\pi C_S R_S}$, and $f_p = \dfrac{1 + g_{fs} R_S}{2\pi C_S R_S}$ (21-38)

If $f_p >> f_o$, then $f_1 \approx f_p$. (21-39)

Gain-Bandwidth Product

For the first stage, $|A_{v(MF)}| f_2 = \dfrac{g_{fs}}{2\pi C_i}$ (21-40)

Common-Drain Amplifier or Source Follower High Frequency Response

At high frequencies, if $g_{fs} R_S >> 1$

$$A_{v(HF)} \simeq \frac{g_{fs} + j\omega C_{gs}}{g_{fs} + j\omega C_T}. \qquad (21\text{-}41)$$

$$C_T = C_{gs} + C_{ds} + C_{sn}, \qquad (21\text{-}42)$$

where C_{sn} is the source-to-ground capacitance.

$$f_2 \approx \frac{g_{fs}}{2\pi C_T}. \qquad (21\text{-}43)$$

21-14. OTHER AMPLIFIER CONFIGURATIONS

Fig. 21-20(a) shows two common-source amplifiers in cascade to provide a voltage gain of approximately 400. This is a low-noise amplifier, which explains the presence of the large capacitors (250 μF) on the power supply lead. The 0.6 μF capacitor will reduce any high-frequency noise amplified by the first stage.

The circuit of Fig. 21-20(b) shows how an FET and a bipolar transistor may be interconnected to give an amplifier with an input impedance in excess of 10 M ohms at 1 kHz. Note how the bandwidth decreases with increasing value of generator internal resistance, R_G. This is due to the input capacitance of the 2N3330 causing appreciable loading, which is offset by the negative feedback from the 1 K ohm resistor in the collector lead of the output transistor.

21-15. SUMMARY

(1) The junction field-effect transistor (JFET) is a voltage-controlled device which has a high input resistance (in the order of 100 M ohms), a high degree of isolation between input and output, and a low noise figure.

(2) The main disadvantage of a JFET is its relatively low gain-bandwidth product, due to the junction capacitive effects.

(3) Both N- and P-channel JFETs exist, with negative and positive gate voltages, respectively, controlling the drain current from the source.

(4) The FET got its name from the electric *field* that extends into the channel providing the *effect* of a decrease in conductivity through the *transistor.*

(5) The FET's common-source characteristics may be divided into a channel ohmic region and a pinch-off region.

(6) The three main characteristics of the FET are:

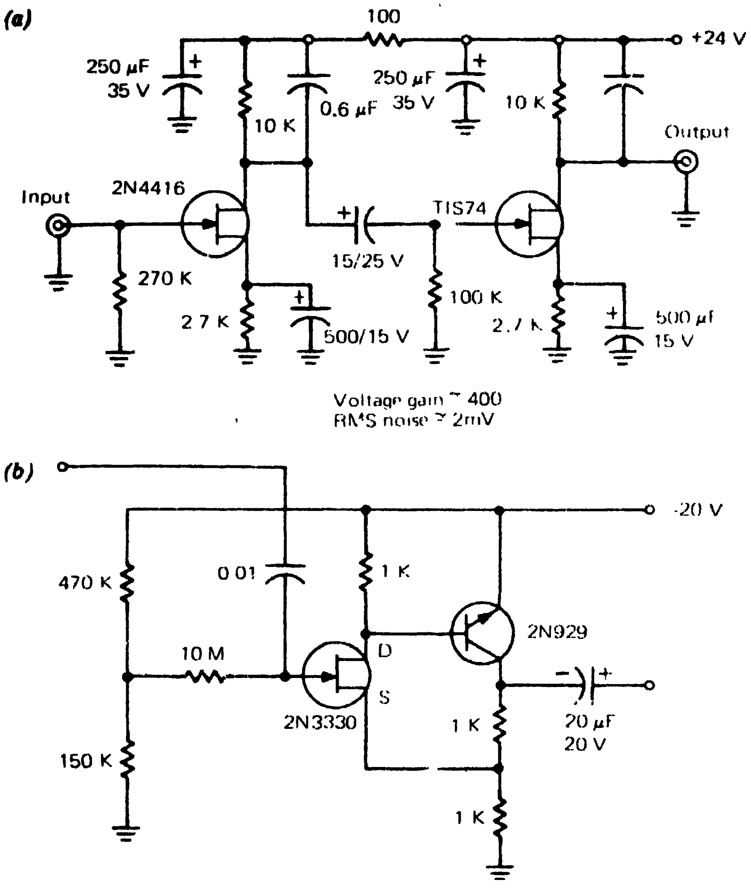
(a) Drain resistance, $r_d = \dfrac{\Delta V_{DS}}{\Delta I_D}$, V_{GS} constant.-

(b) Forward transconductance, $g_{fs} = \dfrac{\Delta I_D}{\Delta V_{GS}}$ V_{DS} constant.

(c) Amplification factor, $\mu = \dfrac{\Delta V_{DS}}{\Delta V_{GS}}$, I_D constant.

$\mu = r_d g_{fs}$.

(7) An increase in temperature reduces g_{fs} and I_D, but increases r_d; gate current, however, increases with temperature, since it consists of


(a)
100
250 μF 35 V
10 K
0.6 μF
250 μF 35 V
10 K
+24 V
Output
Input
2N4416
TIS74
15/25 V
270 K
2.7 K
500/15 V
100 K
2.7 K
500 μF 15 V
(b)
-20 V
470 K
0.01
1 K
2N929
D
10 M
S
2N3330
1 K
20 μF 20 V
150 K
1 K

Voltage gain ≈ 400
RMS noise ≈ 2mV

Voltage gain > 2.0
Distortion < .1% at 3 V pp output
Input impedance @ 1 kHz > 10 M ohms
Response ±1 dB 20 to 20 kHz R_G = 600 ohms
20 to 16 kHz R_G = 1 M ohm
20 to 2.9 kHz R_G = 100 M ohms

Fig. 21-20. (a) Low-noise, high-voltage-gain cascaded amplifier. (b) Direct-coupled FET-transistor amplifier. (Courtesy of Texas Instruments, Incorporated)

minority carriers.

(8) A field-effect transistor may have a combination of self bias and fixed bias, to provide stability of the quiescent drain current against device and temperature variation.

(9) The common-source ac amplifier may be analyzed using load lines drawn on the characteristic curves, the operating point being at the intersection with the bias equation, given by $V_{GS} = -I_D R_S$.

(10) The common-source voltage gain may be

obtained from a small-signal equivalent circuit similar to that for a tube,

$$A_v = \frac{-\mu R_L}{r_d + R_L} = -g_{fs}\frac{R_L}{1 + \frac{R_L}{r_d}}$$

while the input resistance is determined by the gate resistor, R_G, typically 1 M ohm.

(11) The gate-to-drain capacitance, C_{gd}, is mainly responsible for limiting the high frequency operation of the common-source amplifier since the total input capacitance given by the Miller effect is

$$C_i = C_{gs} + (1 - A_v)C_{gd},\ A_v \text{ negative.}$$

(12) The common-drain amplifier or source follower has a voltage gain less than one, but a higher input resistance with a bootstrapped bias circuit. Its output resistance is low.

(13) The common-gate amplifier has a characteristically low input resistance and a high output resistance, although its voltage gain can be quite high. It has few applications.

(14) The equations for the frequency response characteristics of FET amplifiers are summarized in Section 21-13. The higher value of feedback capacitance from drain-to-gate limits the bandwidth compared with tubes or bipolar transistors.

QUESTIONS

Q21-1. (a) What are the two main types of field-effect transistor?

(b) What is the main advantage of the FET over a conventional transistor?

(c) What do the terms "unipolar" and "bipolar" refer to?

Q21-2. (a) What is meant by the "substrate"?

(b) What type of material is the substrate in a P-channel JFET?

(c) What is the normal polarity of the gate with respect to the source in a P-channel JFET?

(d) What are the carriers in a P-channel JFET?

(e) What effect will an increased positive voltage on the gate of a P-channel JFET have on the drain current?

Q21-4. Define:

(a) The pinch-off voltage.
(b) Channel ohmic region.
(c) Drain resistance.
(d) Transconductance.
(e) I_{DSS}
(f) I_{GSS}

Q21-5. (a) What effect does an increase in temperature have on drain current over most of the transistor's characteristics? What is this due to?

(b) What happens to the drain current at large gate-to-source reverse voltages as temperature increases? Why?

(c) What disadvantage arises from biasing the transistor for zero drain current drift with temperature variation?

(d) What polarity of temperature coefficient does the reverse gate current have? Why?

Q21-6. Consider the common-source ac amplifier in Fig. 21-10.

(a) What configuration does this correspond to in bipolar transistor amplifiers? What is the major difference?

(b) As R_L is increased, what does the voltage gain approach?

(c) What is the input resistance seen by v_s determined by?

(d) What is R_G used for? What limits the size of R_G?

Q21-7. (a) Which capacitance is the most troublesome in an FET amplifier?

(b) How does the total input capacitance vary with R_L?

Q21-8. (a) What method is used to bias an FET against device and temperature variation? Explain how this is effective.

(b) What do transfer curves for an FET consist of?

(c) What is a bias line? What does it imply if the bias line is not very steep?

Q21-9. (a) What type of FET circuit produces a voltage gain less than one?

(b) How may the bias circuit be modified to produce a higher overall input resistance? What two factors make this successful?

(c) Explain how bootstrapping occurs in Fig. 21-17.

(d) Why is the input capacitance very low in Fig. 21-17?

Q21-10. What characteristics does a common-gate FET amplifier have?

Q21-11. (a) What is the input resistance to the amplifier in Fig. 21-20(a)?

(b) Why is the input resistance much higher in Fig. 21-20(b)?

(c) What type of bias do the 150 K ohm and 470 K ohm resistors provide in Fig. 21-20(b)? What provides the reverse bias?

(d) What type of transistor is the 2N929? Does this not usually require a positive supply voltage? How has this been effectively provided here?

(e) How has forward bias been provided for the 2N929?

PROBLEMS

P21-1. The drain resistance and transconductance of a JFET are to be determined experimentally. With $V_{GS} = 0$ V and $V_{DS} = 7$ V, the drain current is 10 mA. When the drain voltage is increased to 15 V with V_{GS} held constant at 0 V, the drain current is increased to 10.25 mA.

When the gate-to-source voltage is now increased to -0.2 V while V_{DS} is held constant at 15 V, the drain current decreases to 9.65 mA. Determine:

(a) The drain resistance
(b) The transconductance
(c) The amplification factor
(d) The type of transistor.

P21-2. A 2N5245 silicon N-channel FET has $I_{GSS} = 1$ nA at $V_{GS} = -20$ V and $T_A = 25$°C. This current increases to 0.5 μA at $T_A = 100$°C. Determine the maximum and minimum values of input resistance between gate and source over a temperature range from 25°C to 100°C.

P21-3. A common-source amplifier uses self bias, with the result that the gate-to-source bias voltage may vary from 0.75 V to 2.3 V as the temperature varies from 125°C to -55°C, depending upon the transistor used. Use the curves in Fig. 21-9 to determine the extremes of drain current that would result and the corresponding values of transconductance over this temperature range.

P21-4. Assume the common-source amplifier of Fig. 21-10 employs a 4012 P-channel JFET, that obeys the curves in Fig. 21-9. $R_L = 10$ K ohms, $R_S = 5.1$ K ohms, $R_G = 1$ M ohm, $V_{DD} = -30$ V. If the quiescent drain current at 25°C is 1 mA and its drain resistance, $r_d = 100$ K ohms, determine:

(a) The minimum and typical voltage gains the amplifier could have at 25°C.
(b) The output resistance seen by v_o.
(c) The input resistance, R_i, if $I_{GSS} = 2$ nA at 25°C.
(d) The overall input resistance, R_i'.
(e) The total input capacitance, C_i, for the typical transistor, taking into account the Miller effect. Given: $C_{iss} = 7$ pF, C_r pF.

P21-5. Repeat P21-4 with: $R_L = 2.7$ K ohms, $R_S = 560$ ohms, $R_G = 1$ M ohm, $V_{DD} = -30$ V;

I_D = 5 mA and r_d = 40 K ohms.

P21-6. In an attempt to bias a 4012 P-channel FET amplifier at a quiescent drain current of 6 mA a 220-ohm self-bias resistor, R_S, is used.

(a) Using the curves in Fig. 21-9(a), determine the extremes of drain current that could result from device and temperature variation, from -55°C to +125°C. (Hint: Draw a bias line of slope $1/R_S$ through the origin.)

(b) Modify the bias arrangement by specifying R_S and the ratio of fixed-bias resistors, R_1 to R_2, to provide I_D = 6 mA ±2 mA over the same device and temperature variation as in (a). V_{DD} = -30 V.

(c) Repeat (b) to provide I_D = 6 mA ±1 mA. (Extend the curve for 125°C, Min, to obtain V_{GS} at 5 mA. This will be a negative value.)

P21-7. In an attempt to bias a 2N5245 N-channel JFET amplifier at a quiescent current of 5 mA at 25°C, a 200-ohm self-bias resistor, R_S, is used. The transistor may have an I_{DSS} anywhere between 5 mA and 15 mA, as shown in the transfer curves below, which are drawn for 25°C.

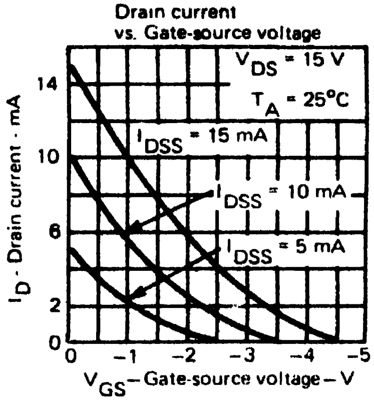
(a) Determine the extreme values of drain current that could result with this type of bias at 25°C.

(b) Determine the new value of R_S and the ratio of values of R_1 to R_2 to keep I_D within 1 mA of the desired value of 5 mA at 25°C. V_{DD} = 20 V.

(c) Suggest values for R_1 and R_2 if I_{GSS} = 1 nA at 25°C and operation is to remain at 25°C.

(d) Repeat (b) for I_D = 5 mA ±0.5 mA.

P21-8. (a) What value of self-bias resistor, R_S, (without any fixed bias) could maintain the quiescent drain current of a 4012 FET between 1 mA and 2 mA over a temperature variation of -55°C to +125°C? Use the curves in Fig. 21-9 and take into account the minimum and maximum device variations that could occur.



Transfer curves of the 2N5245 N-channel FET for P21-7. (Courtesy of Texas Instruments, Incorporated)

(b) If this same resistor were used with a fixed-bias forward voltage of -6 V (due to R_1 and R_2), what would be the extremes of drain current possible?

(c) Suggest values to use for R_1 and R_2 in (b) if V_{DD} = -18 V and I_{GSS} = 1 μA at 100°C.

P21-9. A common-drain amplifier, as in Fig. 21-16, uses a 4012 FET. When R_S = 2.2 K ohms, I_D = 2 mA and r_d = 80 K ohms. R_G = 2.2 M ohms.

(a) Use the curve in Fig. 21-9(b) to find the typical voltage gain.

(b) What is the output resistance, R_o'?

(c) What is the input resistance, R_i'?

(d) The bias circuit is modified to that in Fig. 21-17, with R_A = 220 ohms, R_B = 2.2 K ohms, R_G = 2.2 M ohms. The quiescent drain current is now 5 mA, at which r_d = 40 K ohms. Repeat (a), (b), and (c).

(e) Find the total input capacitance for the conditions in (d), given: C_{iss} = 7 pF, C_{rss} = 2 pF.

P21-10. It is required that a common-drain amplifier, having the configuration shown in Fig. 21-17, have an input resistance of 10 M ohms and an output resistance of 200 ohms. An FET with g_{fs} = 4500 μmhos and r_d = 100 K ohms is to be used.

(a) Determine the necessary value for R_S.
(b) Obtain the voltage gain.
(c) Determine the values of R_A and R_B if R_G 2.2 M ohms.

P21-11. Determine the voltage gain, input resistance, and output resistance of the common-gate amplifier in Fig. 21-18 if r_d = 100 K ohms, g_{fs} = 3000 μmhos, R_L = 10 K ohms, R_s = 600 ohms.

P21-12. Two identical FET common-source amplifiers are cascaded as in Fig. 21-19. R_L = 10 K ohms, R_S = 2.2 K ohms, C_S = 50 μF, R_G = 1 M ohm, C_c = 1 μF, g_{fs} = 5000 μmhos, r_d = 50 K ohms, C_{ds} = 2 pF, C_{gs} = 8 pF, C_{gd} = 4 pF, C_{stray} = 20 pF. Calculate:

(a) The mid-frequency voltage gain of each stage.
(b) The upper 3-dB frequency, f_2, for the first stage, V_o/V_s, and for the second stage, V_1/V_o.
(c) The gain and phase angle of V_o/V_s at 100 kHz.
(d) The lower 3-dB frequency, f_1, for the first stage, V_o/V_s, and for the second stage, V_1/V_o.
(e) The gain and phase angle of V_o/V_s at 10 Hz.
(f) The overall bandwidth of the amplifier (estimate).
(g) The gain-bandwidth product of the first stage.

22

MOSFETs AND OTHER APPLICATIONS OF FETs

22-0. INTRODUCTION

In the last chapter, we examined the junction field-effect transistor or JFET, of which there is only a depletion type. We now look at the other family, the MOSFET, in which, as the FET "tree" in Fig. 21-1 shows, there are two types. These are the depletion and enhancement types. Also known as the insulated gate FET or IGFET, the MOSFET derives its name from the fact that its metal gate is insulated by a very thin oxide layer from the semiconductor channel. (Note the underlined letters.)

In this chapter, we also look at the variable drain resistance property of a field-effect transistor and how this may be used in applications other than amplifiers. The important development of the dual-gate MOSFET is examined and how it is particularly suited for automatic gain control circuits and mixers. Finally, the field-effect diode or current regulating diode (CRD) is found to approximate a very good current source and has some unique applications.

22-1. THE DEPLETION MOSFET

As shown by the diagram in Fig. 22-1, there is no PN junction between gate and channel but rather a capacitor, consisting of the metal gate contact, a dielectric of silicon dioxide, and the channel. It is this construction which accounts for the very large input resistance of 10^{10} to 10^{15} ohms and is the major difference from the JFET.

The symbol clearly shows the gate insulated from the channel. Some manufacturers internally connect the bulk to the source, but there are times when a circuit requires the two to be separated. An N-channel MOSFET is shown in Fig. 22-1. The symbol clearly shows where the PN junction is located. (A P-channel would have the arrow reversed.)

The drain characteristics and transfer curve for the depletion MOSFET are shown in Fig. 22-2.

When $V_{GS} = 0$ V, a significant current flows for a given V_{DS} (like the JFET). When the gate (one plate of the capacitor) is made negative, the other plate, the channel, has a positive charge induced in it opposite the gate. This serves to deplete the channel of majority carriers (electrons) so the conductivity decreases, giving rise to characteristic curves like the JFET. However, unlike the JFET, there is no PN junction to keep reverse biased, so the gate is allowed to be made positive. Under this condition, a negative charge is induced in the channel, thereby increasing carriers and *enhancing* current flow. Thus, this device can be operated in depletion and enhancement modes although it is termed a depletion type to differentiate it from the second MOSFET type that can only work in the enhancement mode. (A JFET is inherently a depletion type, but since no enhancement JFET is available no distinction is necessary.)

The depletion MOSFET can thus be operated at zero bias in an amplifier with no change in input resistance on the positive and negative excursions of the input signal voltage.

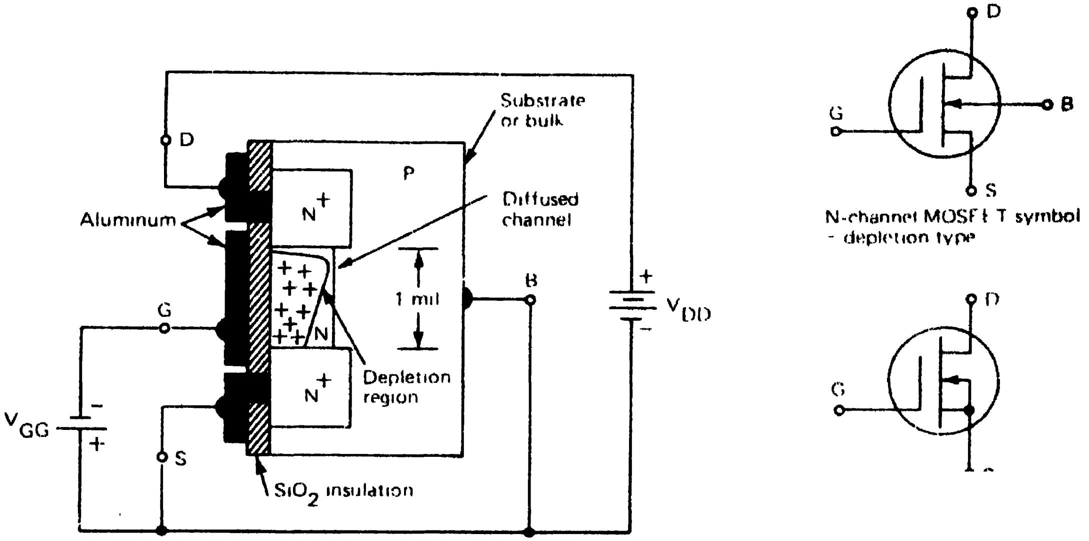

D
Substrate or bulk
P
Aluminum
N+
Diffused channel
1 mil
B
+
V_DD
−
G
N
Depletion region
N+
V_GG
−
+
S
SiO2 insulation
D
G
B
S
N-channel MOSFET symbol - depletion type
D
G
S

Fig. 22-1. Depletion type of MOSFET and alternative symbols for N-channel. (Courtesy of Motorola Semiconductor Products Inc.)

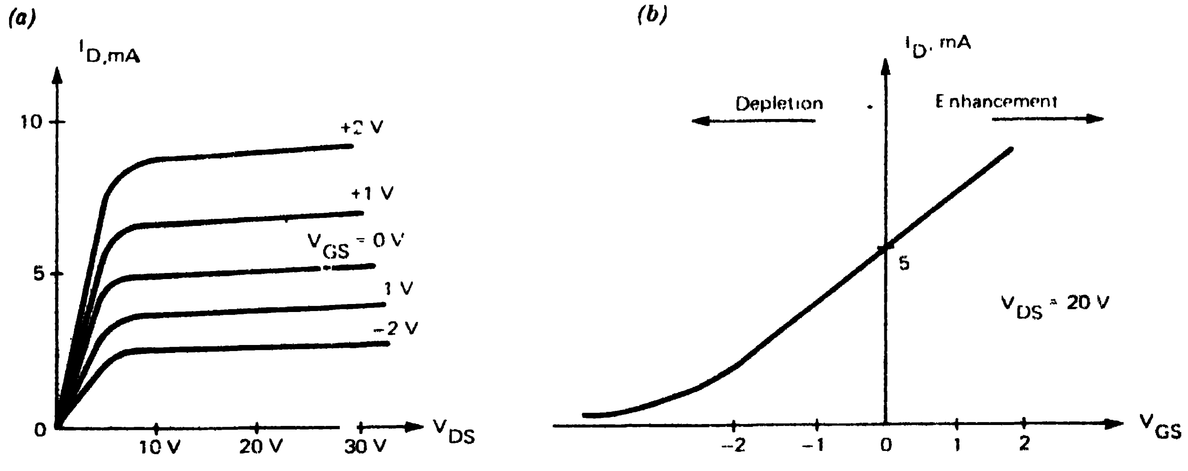


Fig. 22-2. (a) Drain characteristics for a depletion-type MOSFET. (b) Transfer curve for a depletion MOSFET showing how depletion and enhancement operation is possible.

22-2. THE ENHANCEMENT MOSFET

The enhancement MOSFET differs from the depletion type in that no continuous channel exists between source and drain, with the result that no current flows at zero gate voltage. The symbol's "broken channel" represents this condition, as can be seen in Fig. 22-3.

The only way in which current can flow from source to drain is for a positive voltage on the gate to *induce* a channel by drawing the minority carrier electrons in the P-type bulk into a concentrated layer. As can be seen by the transfer characteristics in Fig. 22-4, a minimum or threshold gate voltage is required before drain current will flow.

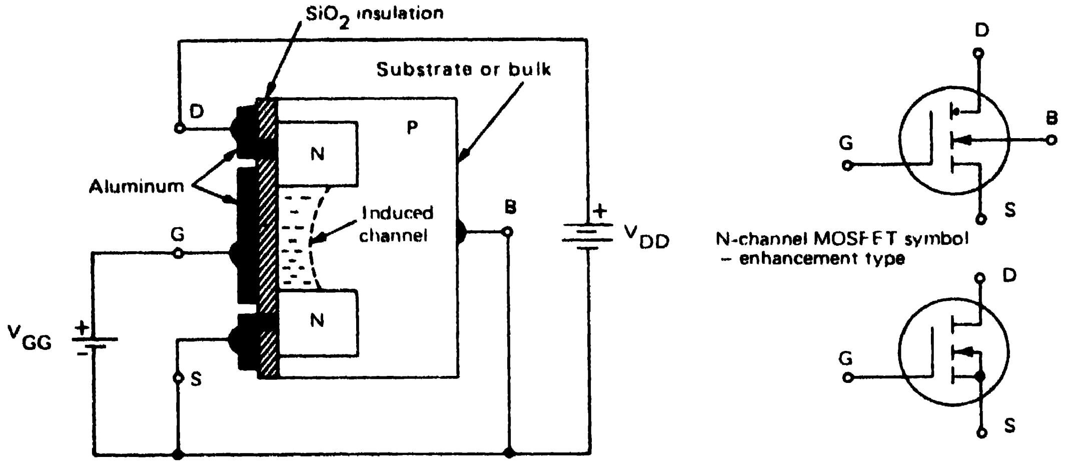

SiO_2 insulation
Substrate or bulk
D
P
N
Aluminum
Induced channel
B
G
V_{DD}
V_{GG}
N
S
D
B
G
S
N-channel MOSFET symbol
– enhancement type
D
G
S

Fig. 22-3. Enhancement type of MOSFET with positive gate voltage inducing a channel. (Courtesy of Motorola Semiconductor Products Inc.)

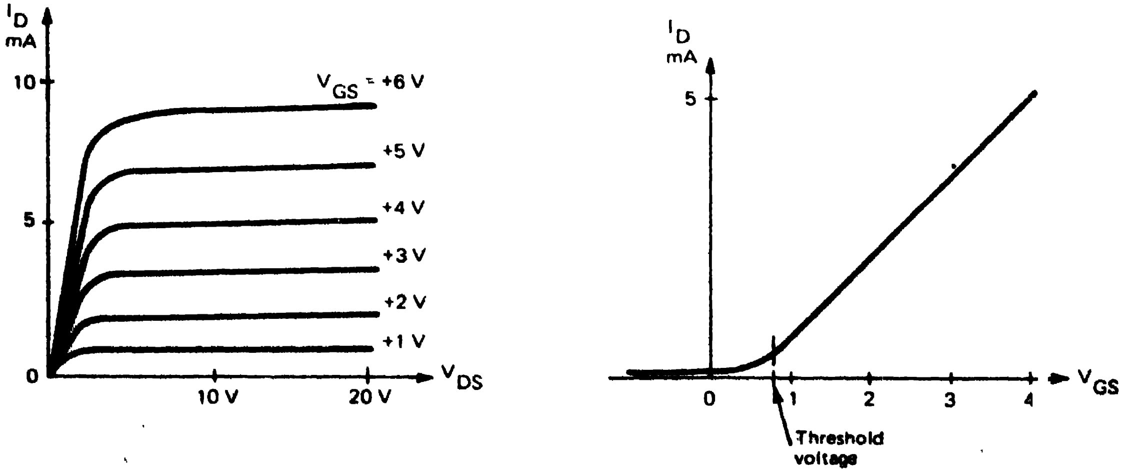


Fig. 22-4. Typical drain characteristics and transfer curve for an N-channel enhancement type of MOSFET.

This type of FET is very useful in switching applications, since no gate voltage is required to hold it off, and when the device is to be switched on the application of the same polarity as the drain will achieve this condition. Also, the threshold voltage provides some "noise immunity" to prevent false switching.

22-3. DIFFERENCES BETWEEN JFETs AND MOSFETs

The principal difference between the JFET and MOSFET is the much higher input resistance of the latter because there is a very small gate leakage current. The characteristic curves of the JFET are flatter than those of the MOSFET, indicating a higher drain resistance. Capacitive effects may be considerably lower in dual gate MOSFETs, as explained in Section 22-5. Comparisons are shown in Table 22-1.

TABLE 22-1 COMPARISON OF JFETs AND MOSFETs

	JFET	MOSFET
g_{fs}	1000 to 25,000 μmhos	1000 to 20,000 μmhos
r_d	0.1 to 1 M ohm	1 to 50 K ohms
I_{GSS}	0.1 to 10 nA	0.1 to 10 pA
r_{gs}	$>10^9$ ohms	$>10^{13}$ ohms
C_{rss}	1 to 4 pF	0.005 to 1 pF

The MOSFET may be used in any of the circuits covered for the JFET, with the bulk or substrate connected to the source. Thus, all the equations apply equally well to the MOSFET and JFET in amplifier connections.

22-4. HANDLING PRECAUTIONS FOR MOSFETs

The very thin layer of SiO_2, silicon dioxide, between gate and channel is very susceptible to high voltages and is very easily punctured, even by static electricity resulting from the transistor sliding around in a plastic bag. A large electrostatic discharge may also result from a person who picks up the transistor by its case and brushes the gate lead against some grounded object. In a relatively dry atmosphere, a static potential of 300 volts is not uncommon on a person who has high resistance soles on his shoes.

As a result, MOSFETs are protected by a shorting ring that is wrapped around all four terminals during shipping and must remain in place until after the device is soldered into position. See Fig. 22-5. Prior to soldering, the technician should use a shorting strap to discharge his static electricity and make sure the tip of the soldering iron is grounded. Once in the circuit, there are usually low enough resistances present to prevent any excessive accumulation of electrostatic charge. However, the MOSFET should never be inserted into or removed from a circuit with the power on. The JFET is not subject to these restrictions, and even some MOSFETs have a built-in gate protection as explained in Section 22-6.

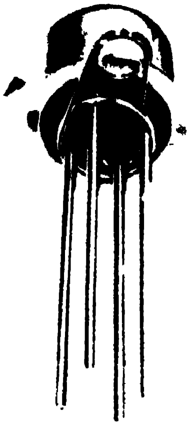


Fig. 22-5. Illustration shows shorting spring for MOS field-effect transistors that do not contain the integral-gate protection. (Courtesy of RCA Corp.)

22-5. DUAL-GATE MOSFETs

Apart from the higher input resistance, a MOSFET is not much different from a JFET in its characteristics. Even the capacitive effect from drain-to-gate in a single-gate MOSFET is of the same order of magnitude as for the JFET. But a most significant advantage derives from the MOSFET construction when a second insulated gate is provided. A cross-sectional view of a dual-gate N-channel depletion-type MOSFET is shown in Fig. 22-6, along with its symbol.

The device acts as if it were two FETs in series, with the N-type "block" in the middle acting as a drain for Unit Number 1 and as a source for Unit Number 2. Consequently the flow of current through the device is controlled by the voltage of both gates. That is, either gate when made sufficiently negative (for an N-channel) can cut off the drain current. However, it is the simultaneous control of current by voltages on both gates that gives this device important applications, such as a gain-controlled amplifier, mixer, or demodulator.

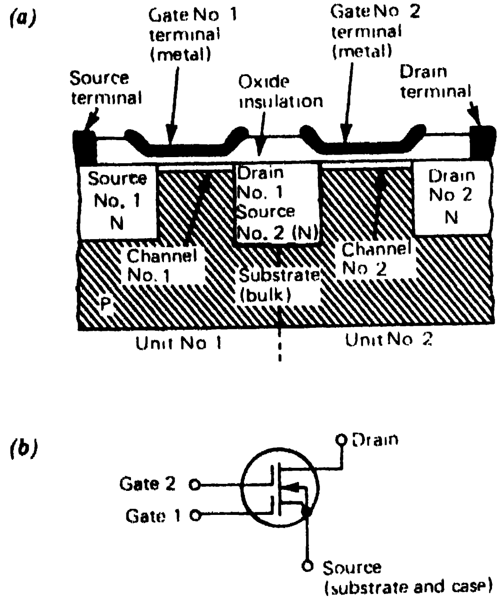


Fig. 22-6. (a) Cross-sectional view of a dual-gate N-channel depletion-type MOSFET. (b) Symbol showing the substrate and case tied to the source. (Courtesy of RCA Corp.)

Characteristics of a Dual-Gate MOSFET

The dual-gate MOSFET may be considered to be the counterpart of a tetrode (or pentode), due to the controlling capabilities exerted by the two gates. And, although Fig. 22-6(a) may suggest a symmetrical device, it is Gate Number 1 that provides the higher forward transconductance (g_{fs}) of the two units. (Each gate has the same value of cutoff voltage, however.) Therefore, the signal to be amplified is connected to Gate 1, while Gate 2 has the voltage that is to control the gain (in an automatic gain control (agc) application).

Typical drain transfer characteristics of an N-channel depletion-type, dual-gate MOSFET are shown in Fig. 22-7, for a 3N200.

Consider the Gate 1-to-source voltage, V_{G1S}, to be 0.5 V. The drain current will be approximately 2 mA if $V_{G2S} = 0$ V, but will increase to 11 mA if $V_{G2S} = 4$ V. This effect may be represented by the simple equivalent circuit shown in Fig. 22-7(c), where R_2 represents the variation that can be obtained through control of Unit 2. This control may also be seen by referring to the manner in which the forward transconductance of Unit 1 varies with V_{G2S}. This is illustrated in Fig. 22-7(b), where, when $V_{G1S} = 0.5$ V, $g_{fs} = 1500\ \mu$mhos @ $V_{G2S} = 0$ V. When $V_{G2S} = 4$ V, g_{fs} increases to 15,000 μmhos.

Another reason for using Gate No. 1 as the signal input is to provide a higher degree of isolation (and shielding) from the drain. In this way, the reverse transfer capacitance C_{rss} (C_{gd}) may be made as small as 0.005 to 0.03 pF from drain to Gate No. 1. This value is typical for the 3N200 shown in Fig. 22-7.

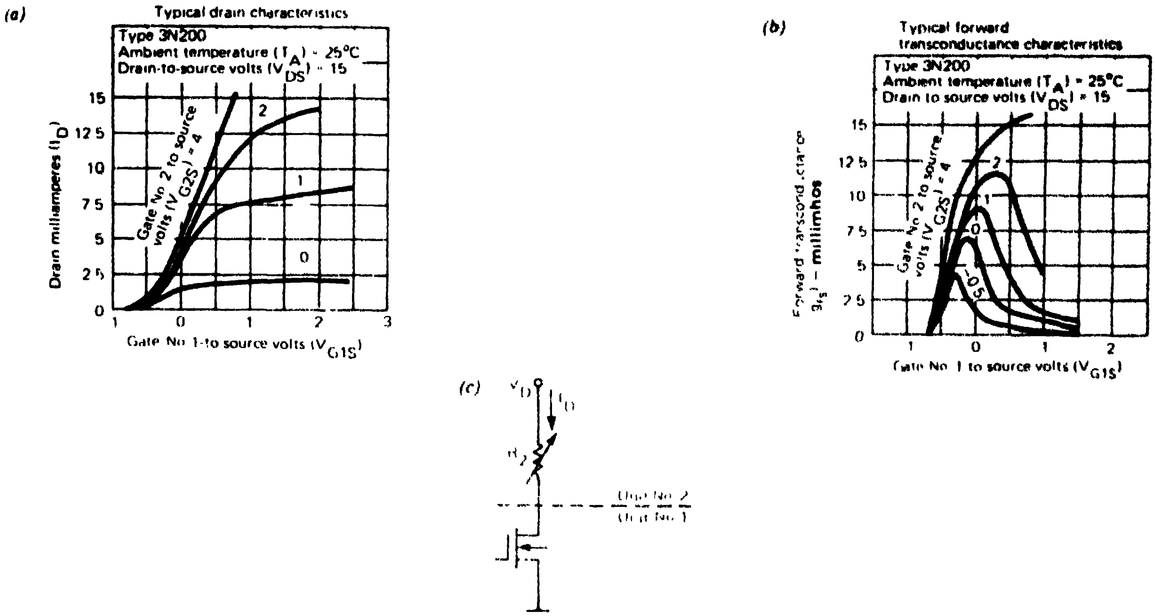


Fig. 22-7. (a) Typical drain transfer characteristics of a 3N200 dual-gate N-channel depletion type MOSFET. (b) Showing how the transconductance of Unit 1 varies with V_{G1S} and depends upon the Gate 2-to-source volts, V_{G2S} of Unit 2. (c) Simple equivalent circuit of the dual-gate N-channel depletion type of MOSFET. (Courtesy of RCA Corp.)

Consequently, this device is used in applications up to 500 MHz, a frequency much higher than is possible with the JFET. Evidently, the second gate in the MOSFET has done for the FET what the screen grid (and suppressor grid) did for the triode modified into a pentode. The single-gate MOSFET has a C_{rss} in the order of 1 pF, like that of a JFET.

Before looking at applications of the dual-gate MOSFET let us look at a development that made MOSFETs as acceptable as any other solid-state device.

22-6. INTEGRAL GATE PROTECTION

The handling precautions for MOSFETs listed in Section 22-4 make the MOSFET an awkward device to handle and set it aside from other solid state devices as far as advantages are concerned. The advent of integral gate-protection, a system built into the device to get around the problem of high voltage on the gate causing a puncturing of the oxide layer, removed the need for the listed precautions. The manner in which this was done is shown in the cross-sectional view of Fig. 22-8(a). The symbol clearly shows that between each gate and the source is placed a back-to-back (or front-to-front) pair of diodes, which are built right into the P-type substrate (one pair at each side of the MOSFET proper).

These diodes are designed so that if either gate exceeds +10 V typically, with respect to the source, the upper diode will conduct and the lower diode will break down (Zener effect), providing a shunt path for excessive charge from gate to source. Likewise, if the gate

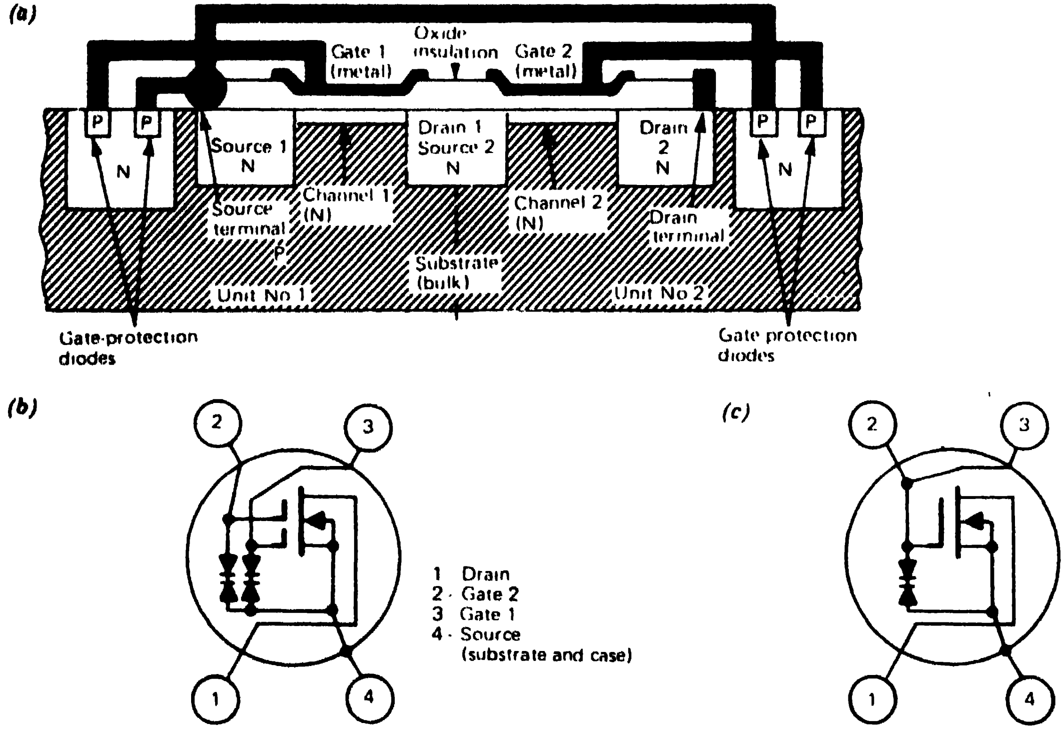

(a)
Oxide insulation
Gate 1 (metal)
Gate 2 (metal)
P
P
N
Source 1 N
Drain 1 Source 2 N
Drain 2 N
P
P
N
Channel 1 (N)
Source terminal
P
Channel 2 (N)
Drain terminal
Substrate (bulk)
Unit No 1
Unit No 2
Gate-protection diodes
Gate protection diodes
(b)
2
3
1
4
1 Drain
2 - Gate 2
3 Gate 1
4 - Source (substrate and case)
(c)
2
3
1
4

Fig. 22-8. (a) Dual-gate-protected N-channel depletion-type MOSFET, cross-sectional view. (b) Symbol of the dual-gate-protected MOSFET. (c) Showing how a dual-gate may be operated as a single-gate-protected MOSFET by joining Gates 1 and 2. (Courtesy of RCA Corp.)

voltage exceeds -10 V, the lower diode conducts and the upper diode breaks down. For the 3N200 referred to earlier, the breakdown voltage from either gate to source is between +6.5 and +13 V or -6.5 and -13 V. The normal application of signal or control voltages between ± 6.5 V will not be affected by the gate-protection diodes. They only clip voltages in excess of these values of $V_{(BR)G1SSF}$ and $V_{(BR)G1SSR}$.

If the dual-gate-protected MOSFET is to be operated as a single gate-protected device, it is only necessary to interconnect the two gates as shown in Fig. 22-8(c). One minor disadvantage of the integral gate-protection is the increase in gate leakage current, I_{G1SS}, and I_{G2SS} to 50 nA at 25°C in the 3N200. However, the increased capacitive effect is negligible. In addition to providing protection against preinstallation high static voltages, these diodes also guard against in-circuit transients.

22-7. TESTING FIELD-EFFECT TRANSISTORS

A method of testing field-effect transistors, whether JFET or MOSFET, involves checking to see if the appropriate voltage on the gate causes the device to conduct current from source to drain. A simple ohmmeter check may not be used with MOSFETs, although a check of the PN junctions from gate to source or drain will give some indication of correct operation in a JFET.

A go/no-go test circuit that is appropriate for N-channel depletion types (JFET or MOSFET) and P-channel enhancement types (MOSFET) is shown in Fig. 22-9. For an N-channel depletion type when the switch is open, sufficient drain current should flow through the 270-ohm resistor to forward bias the bipolar transistor. This will switch on the transistor, and the lamp should light. If the switch is now closed, the FET should be biased off, and the light should go off. If the lamp lights no matter which position the switch is in, the FET is shorted. If the lamp does not light in either position of the switch the FET is open. For a P-channel enhancement type, the reverse indications of the lamp would be observed.

By changing the polarity of the gate supply, other transistor types may be tested (see student problem).

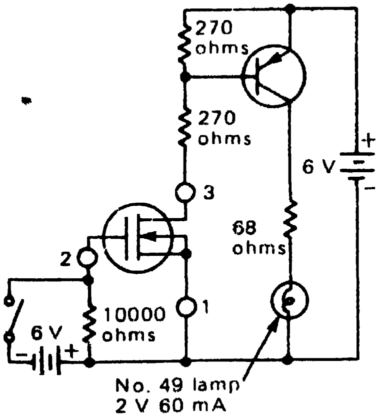


Fig. 22-9. Go/no-go test circuit for field-effect transistors of the N-channel depletion and the P-channel enhancement types. (Courtesy of RCA Corp.)

22-8. APPLICATION OF A DUAL-GATE MOSFET IN AN AGC AMPLIFIER

Shown in Fig. 22-10 is a dual-gate N-channel depletion-type MOSFET amplifier with automatic gain control (AGC). This type of amplifier is usually found in RF (radio frequncy) applications, such as radio and TV. V_{AGC} is a positive voltage fed back from later stages, whose magnitude varies inversely with the strength of signal. (It is often referred to as reverse AGC.) Should the radio or TV be tuned to a stronger station, tending to increase the audio output level, a smaller positive voltage or even a negative voltage will be fed back to Gate 2, reducing the gain of the amplifier and automatically maintaining the volume at the preset level.

Example 22-1. Using the characteristics of the 3N200 shown in Fig. 22-7, determine:

(a) The values of bias resistors R_1 and R_2, and the drain supply voltage, V_{DD}, to operate the transistor at a drain-to-source voltage of 15 V at a forward transconductance $g_{fs} = 15{,}000\ \mu$mhos.

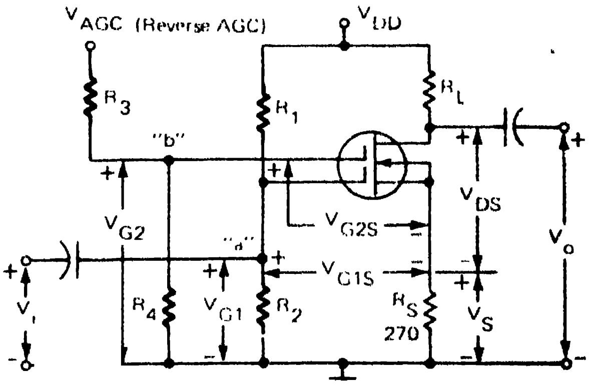


Fig. 22-10. Dual-gate N-channel MOSFET amplifier with AGC.

(b) The values of bias resistors R_3 and R_4 to provide the conditions in (a) when $V_{AGC} = +8$ V.

(c) The maximum voltage gain of the amplifier with $V_{AGC} = +8$ V; effective $R_L = 1$ K ohm.

(d) The voltage gain of the amplifier with $V_{AGC} = 0$ V.

Solution

(a) From Fig. 22-7(b), $g_{fs} = 15{,}000\ \mu$mhos when $V_{G1S} = +0.5$ V and $V_{G2S} = +4$ V. From Fig. 22-7(a), for $V_{G1S} = 0.5$ V and $V_{G2S} = 4$ V, $I_D = 11.2$ mA.

Note that R_S has been given as 270 ohms. This has been determined empirically to give the most suitable self-bias for dual-gate FETs.

$$V_L = I_D R_L = 11.2 \times 10^{-3} \times 1000 = 11.2 \text{ V}.$$

$$V_S = I_D R_S = 11.2 \times 10^{-3} \times 270 = 3 \text{ V}.$$

$$V_{G1} = V_{G1S} + V_S = 0.5 + 3 = 3.5 \text{ V}.$$

$$V_{DD} = V_{DS} + V_S + V_L = 15 + 3 + 11.2 \approx 30 \text{ V}.$$

$$V_{G1} = V_{DD} \times \frac{R_2}{R_1 + R_2},$$

$$\text{or } \frac{R_1 + R_2}{R_2} = \frac{V_{DD}}{V_{G1}} = \frac{30}{3.5} = 8.57.$$

To solve explicitly for R_1 and R_2, we must know how much loading on the input, V_i, can be tolerated. This amounts to

$$\frac{R_1 \times R_2}{R_1 + R_2}.$$

Assuming this may be 100 K ohms, we have:

$$(1)\quad \frac{R_1 R_2}{R_1 + R_2} = 100 \text{ K ohms};$$

$$(2)\quad \frac{R_1 + R_2}{R_2} = 8.57$$

Solving these equations we obtain $R_1 = 857$ K ohms and $R_2 = 113$ K ohms. Nearest standard values of 910 K ohms and 120 K ohms could be used.

(b) $V_{G2} = V_{G2S} + V_S = 4 + 3 = 7$ V.

$$V_{G2} = V_{AGC} \times \left(\frac{R_4}{R_3 + R_4}\right),$$

$$\text{or } \frac{R_3 + R_4}{R_4} = \frac{V_{AGC}}{V_{G2}} = \frac{8}{7},$$

from which $R_4 = 7R_3$.

If $R_3 = 10$ K ohms, then $R_4 = 70$ K ohms (or 68 K ohms for nearest standard value). If this represents too heavy a load on the AGC signal, then values could be increased using the given ratio. Note however, that with $V_{AGC} = +8$ V, the control current necessary is only 100 μA.

(c) With $V_{AGC} = +8$ V, $g_{fs} = 15{,}000\ \mu$mhos, effective load $R_L = 1$ K ohm. $A_v \approx -g_{fs} R_L = -15{,}000 \times 10^{-6} \times 1 \times 10^3 = -15$.

(d) With $V_{AGC} = 0$ V, $V_{G2} = 0$ V.

$$\therefore V_{G2S} = -V_S = -I_D R_S = -I_D \times 270.$$

The values of V_{G2S} and I_D may be found by trial and error, since each depends upon the other.

Estimate from Fig. 22-7(a) a curve for $V_{G2S} = -0.2$ V. Then, at $V_{G1S} = 0.5$ V, I_D would be approximately 1 mA. Then $V_{G2S} = -1 \times 10^{-3} \times 270 = -0.27$ V, close to the value assumed. (No greater accuracy can be expected from these curves, and all we want to get is an idea of the controlling action.)

From Fig. 22-7(b), for $V_{G1S} = 0.5$ V and $V_{G2S} = -0.2$ V (estimate a curve between 0 and -0.5 V), $g_{fs} = 1000\ \mu$mhos.

$$A_v \approx -g_{fs}R_L = -1000 \times 10^{-6} \times 1 \times 10^3 = -1.$$

This represents a 24-dB decrease in power gain for an AGC voltage variation from +8 V to 0 V.

A further reduction in gain may be had if V_{AGC} is made sufficiently negative to make V_{G2S} have a value between -0.1 and -3 V, which is the value of the gate-to-source cutoff voltage, $V_{G2S(off)}$.

The circuit shown in Fig. 22-10 is modified when used as an RF amplifier, by coupling the input signal to Gate 1 through a coil to a parallel resonant circuit and circumventing the loading problem that R_1 and R_2 present.

If the circuit is used as a mixer (to mix an RF input signal with a local oscillator to produce sum and difference frequencies at the output), the local oscillator would be connected to point "b," and R_3 connected to V_{DD} instead of the AGC voltage.

22-9. OTHER APPLICATIONS OF FETs

All of the following circuits may use either a JFET or MOSFET, but the JFET has been indicated in many to familiarize those who may be working with these devices in practical circuits, since the JFET is usually easier to handle.

The Phase-Shift Oscillator

Although most of the oscillators shown in Chapter 18 will also work with FETs, the high input resistance of the FET is especially valuable in a phase-shift oscillator, to minimize loading. See Fig. 22-11. As before, the frequency of oscillation is given by

$$f = \frac{1}{2\pi\sqrt{6}RC}.$$

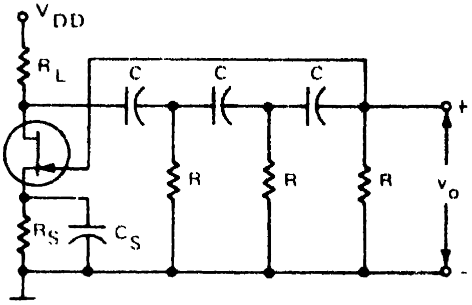


Fig. 22-11. Phase-shift oscillator utilizing an N-channel FET.

Pierce Type of Crystal-Controlled Oscillator

An interesting modification of the Colpitts oscillator (Chapter 18) is the Pierce crystal-controlled oscillator shown in Fig. 22-12.

The way in which feedback is achieved is illustrated in Fig. 22-12(c). The transistor's gate-to-source and drain-to-source capacitances provide a tapped capacitor across the crystal, which acts as a high-Q parallel resonant circuit tuned to some accurately known frequency, determined primarily by the equivalent values of C_p and L in Fig. 22-12(b). Phase reversal from the output at the drain to the input at the gate results from the voltage developed across C_{gs}, with the polarity shown in the diagram.

This type of oscillator, which may have its frequency changed very quickly simply by changing the crystal, is useful in the VHF and UHF range. At lower frequencies, below 2 MHz, the feedback provided by C_{gs} and C_{ds} may be insufficient to maintain oscillations, in which case external capacitors would have to be added to supplement the transistor's capacitance.

Applications of the FET in its Channel Ohmic Region

If the characteristic curves for the FET are examined closely, it is found that in the ohmic region (low values of V_{DS}) each line has a

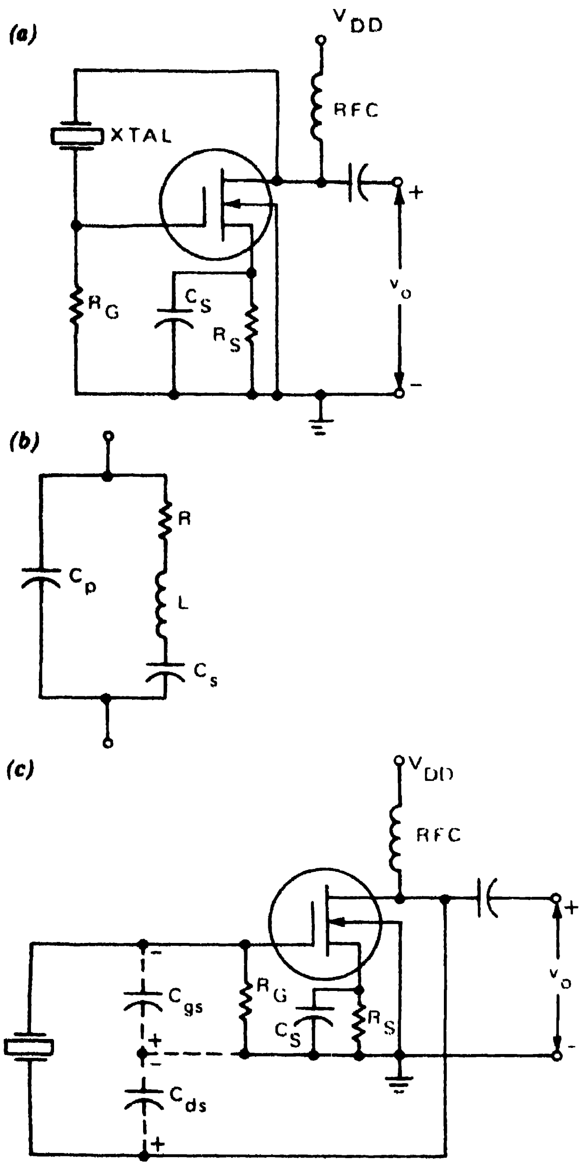


Fig. 22-12. (a) Pierce-type of crystal-controlled oscillator using an N-channel MOSFET. (b) Equivalent circuit of a crystal showing how parallel resonance may occur. (c) Pierce-type oscillator redrawn to show the feedback components provided by the transistor's capacitances.

different slope for each value of gate voltage. See Fig. 22-13(b). Thus, for a given value of V_{DS}, the drain-to-source resistance varies with the gate voltage as depicted in Fig. 22-13(c).

Since R_{DS} $\frac{'DS}{D}$ it is evident that the resistance increases with negative gate voltage, from A to D. The on resistance (at A) may be only a few hundred ohms (evaluated at $V_{GS} = 0$ V and called $R_{DS(on)}$) while the off resistance $R_{DS(off)}$ may be many thousands of megohms. These quantities are very important when the device is used in switching applications.

If the drain-to-source voltage is reversed (negative drain in an N-channel) and the curves obtained for low values of V_{DS}, they will be found to be almost symmetrical with those in the first quadrant.

Thus, if an ac voltage is applied from drain to source (in the order of 0.5 V p-p for example), for a given gate voltage the FET will behave as a linear resistor whose value can be changed or varied by the gate voltage.

Thus, the static resistance $R_{DS} = \frac{V_{DS}}{I_D}$ is also

the dynamic resistance r_d (provided V_{DS} is kept small), because of the straight line relationship between I_D and V_{DS} around the origin.

This voltage variable resistor, VVR, has many applications including a voltage-controlled attenuator, an automatic gain control (AGC) circuit, a switching device in chopper applications to amplify dc voltages, and a solid-state voltmeter with input impedance comparable to that of the VTVM.

We will now examine some of these applications.

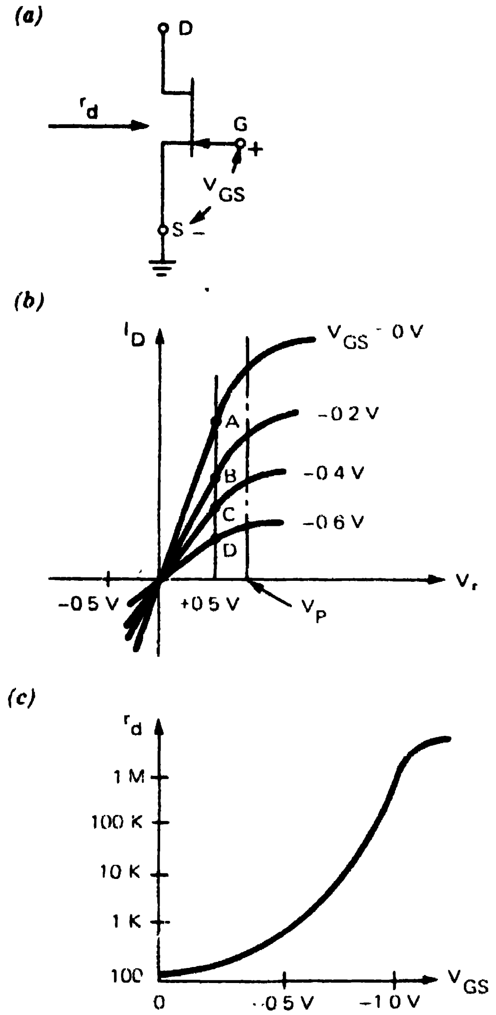

(a)
D
G
S
r_d
V_{GS}
+
−
(b)
I_D
V_{GS} = 0 V
−0.2 V
−0.4 V
−0.6 V
A
B
C
D
−0.5 V
+0.5 V
V_P
V_r
(c)
r_d
1 M
100 K
10 K
1 K
100
0
−0.5 V
−1.0 V
V_{GS}

Fig. 22-13. (a) Showing how the FET may be thought of as a resistance, r_d, between drain and source. (b) Characteristic curves of the FET showing the symmetry about the origin for positive and negative drain voltages, and the linearity of I_D with V_{DS} for a given V_{GS}. (c) Graph showing how the channel resistance, r_d, varies with gate voltage for a fixed V_{DS}.

Application of the FET as a VVR in a Voltage-Controlled Attenuator

Fig. 22-14(a) shows how the FET acts as a variable resistor in a simple voltage divider circuit with R_D. When V_{GS} is small, r_d is small, so that most of V_s appears across R_D. Thus, V_o is near zero.

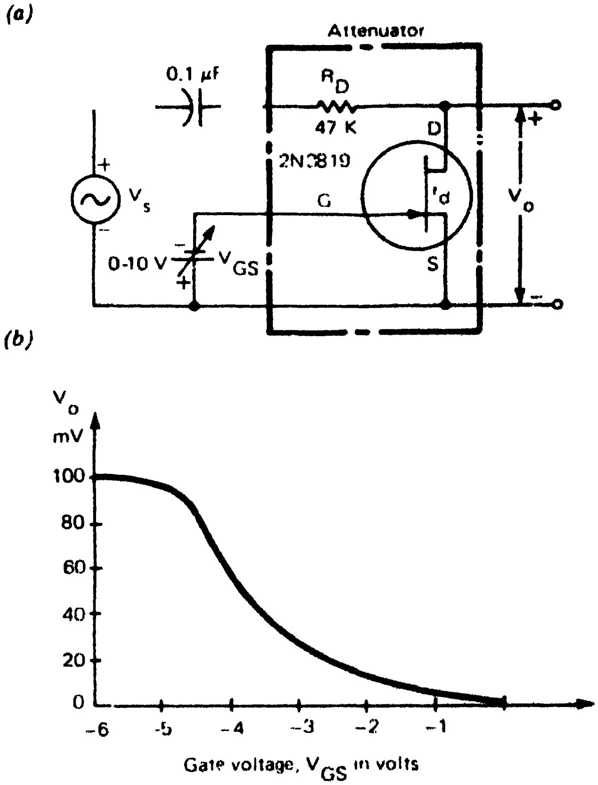

(a)
Attenuator
0.1 μF
R_D
47 K
2N3819
D
G
S
r_d
V_s
0-10 V
V_{GS}
V_o
+
−
(b)
V_o mV
100
80
60
40
20
0
−6
−5
−4
−3
−2
−1
Gate voltage, V_{GS} in volts

Fig. 22-14. (a) Attenuator circuit in which the FET acts as a voltage variable resistor (VVR). (b) Output voltage as a function of gate voltage for the attenuator in (a).

As V_{GS} is made more negative, r_d increases until almost all of V_s appears across the FET, and V_o approaches its maximum value, as shown by the curve in Fig. 22-14(b). It takes a minute amount of power at the gate of the FET to provide a large amount of attenuation with no moving parts. If the input signal V_s is small, typically 100 mV, the distortion introduced by the attenuator will be small — less than two percent.

Application of the FET as a VVR in an AGC Circuit

Another method of providing automatic volume control, AVC, better called AGC, is shown in Fig. 22-15.

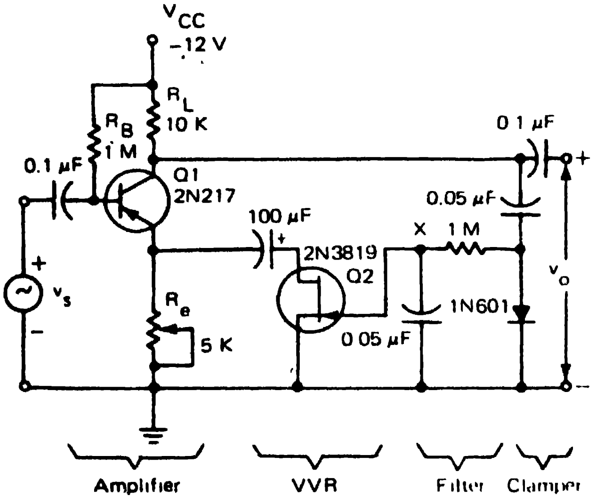


Fig. 22-15. An AGC circuit in which the FET acts as a VVR in parallel with the emitter resistor of the amplifier.

The FET acting as a VVR is in parallel with an unbypassed resistor in the emitter lead of an amplifier. The amount of negative feedback (and thus the gain of the amplifier) is determined by the resistance of the FET (once R_e has been adjusted).

If v_s tends to increase, the increased value of v_o will cause a greater negative dc voltage to appear at X (after rectification and filtering). This will cause r_d of the FET to increase, thus increasing the effective emitter resistance. The resulting increased negative feedback will reduce the gain of Q1 to tend to keep v_o nearly constant. The degree to which this is effective is determined by the sensitivity of the FET's change in resistance with change in v_o, and the extent to which the FET's resistance is comparable to R_e at the desired reference level of v_o.

Application of the FET as a VVR in a Chopper Dc Amplifier

The FET is also useful as a switching device (extreme voltage variable resistor) in chopper applications to amplify dc voltages. Refer to Fig. 22-16.

The negative-going square wave input to the gate of Q1 switches the FET on and off, so that the voltage at X is either zero or equal to V_{in}. This chopped or alternating voltage is amplified in a conventional ac amplifier, rectified and filtered to produce a dc output, V_o, proportional to the input V_{in}. The advantage of this method is that many stages of gain may be provided, capacitively coupled to avoid the bias problems associated with direct coupling in a dc amplifier not utilizing a chopper input.

The higher the chopping frequency of the square-wave input to Q1 the smaller the value of the coupling and bypass capacitors needs to be, as covered in Section 15-5. If V_{in} represents a signal that is varying at some rate, the minimum frequency of chopping (called sampling in some applications) is given by Shannon's sampling theorem:

> "In order to recover the input signal, the sampling frequency should be equal to or larger than twice the frequency of the highest frequency component of the input signal."

FETs used as switches also have application in sampling numerous analog input voltages in a sequential manner, and permitting a

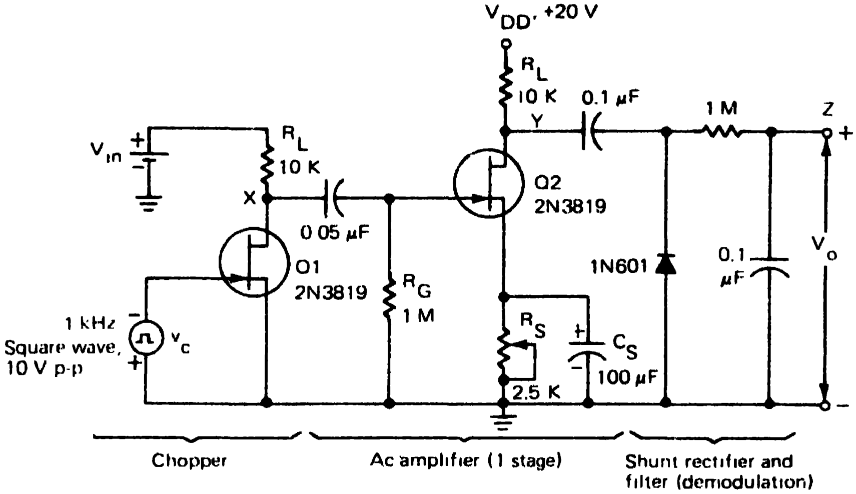


Fig. 22-16. Circuit of a dc amplifier in which Q1 acts as a chopper for V_{in} and Q2 is an ac amplifier. Rectification provides a dc output.

central recording device to give a visual readout of the inputs and how they vary with time.

Application of the FET as a VVR in an FET Voltmeter

Although some solid-state voltmeters use the high input resistance of a field effect transistor at the input to the meter, and then use a conventional transistor bridge circuit for meter deflection, an alternative is shown in Fig. 22-17. Here the FET serves as a high input resistance device and also forms one arm of the bridge. Operation of the circuit should be obvious. If the meter were to be used to measure ac, then some rectification circuit as covered in Chapter 4 would be necessary.

22-10. THE FIELD-EFFECT DIODE

Before leaving the field-effect transistor, let us briefly look at the current-regulating "diode" (CRD), which is essentially an N-channel JFET with an internal short provided between the gate and the source. Since it now has only two active terminals, it is referred to as a diode, although it does not have any rectifying properties.

The principle behind the CRD is that almost constant current flows from drain to source when the drain to source voltage is in excess of the pinch-off voltage and the gate of the JFET is connected to the source. This current has been referred to earlier as I_{DSS}. The almost horizontal line of the CRD's characteristics in Fig. 22-18 indicates how the device may be thought of as a constant-current source provided the voltage is maintained above the "knee" and below the break over voltage, which is typically 100 V. Its equivalent circuit may be represented by either of the two diagrams in Fig. 22-18. I_P is the diode regulator current at a given test voltage, V_T, and may be anywhere between 0.2 mA and 5 mA depending upon the CRD selected. C_T is approximately constant at 8 pF. Z_T is evaluated at V_T and is the dynamic impedance of the diode, with typical values from 25 M ohms for $I_D = 0.2$ mA

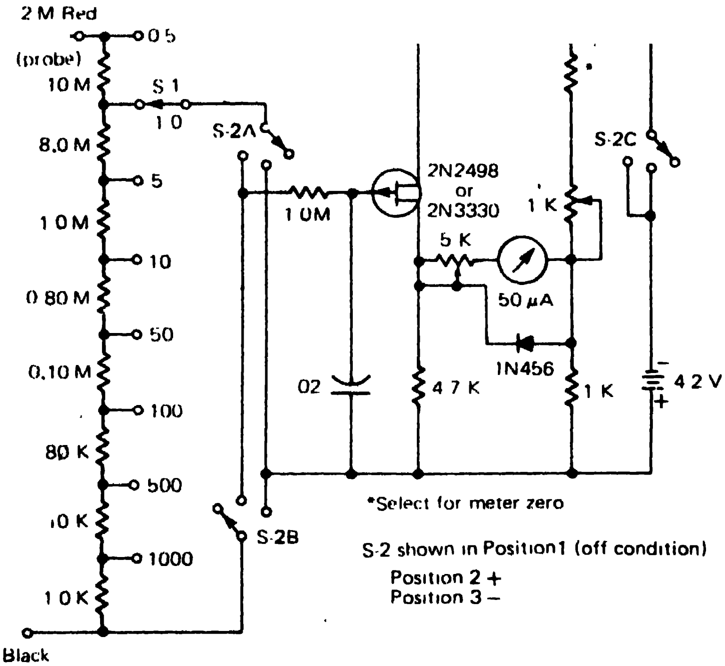

2 M Red
(probe)
0.5
10 M
S 1
1.0
8.0 M
5
1.0 M
10
0.80 M
50
0.10 M
100
80 K
500
10 K
1000
1.0 K
Black
S-2A
S-2B
S-2C
1.0M
2N2498
or
2N3330
5 K
1 K
50 μA
.02
4.7 K
1N456
1 K
4.2 V
*Select for meter zero
S-2 shown in Position1 (off condition)
Position 2 +
Position 3 –

Fig. 22-17. Circuit showing how an FET may be used to provide a high-input-resistance solid-state dc voltmeter. (Courtesy of Texas Instruments, Incorporated)

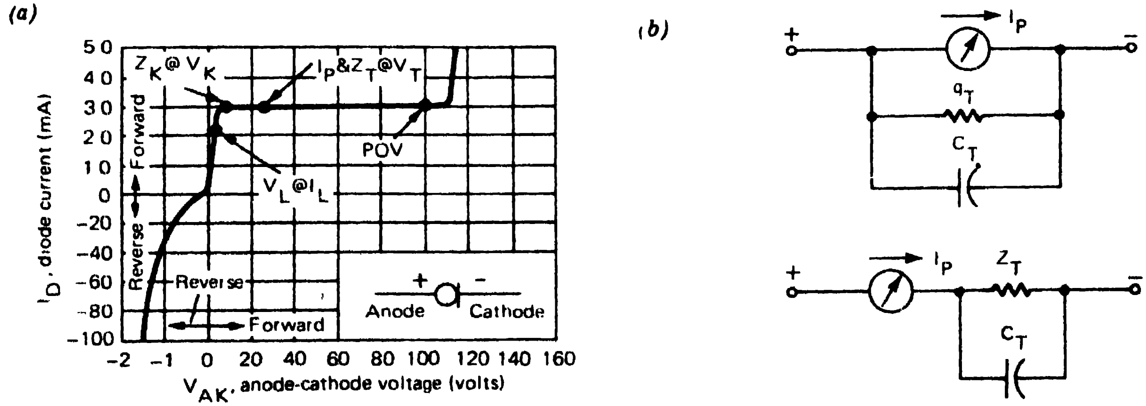


Fig. 22-18. (a) Typical characteristics for a current-regulating diode (CRD). (b) Two equivalent circuits for a CRD, showing a parallel conductance or a series impedance. (Courtesy of Motorola, Inc.)

(1N5283) down to 0.235 M ohm for $I_P = 5$ mA (1N5314).

It is the very high value of Z_T that provides the CRD with its constant current source capability, and, although a JFET operated with its gate at 0 V will show similar characteristics, the construction of the CRD has been especially optimized for high impedance and current-regulating capability.

The current-voltage curve of the CRD should be compared with the Zener diode's characteristics. Whereas the Zener diode provides essentially constant voltage over a wide variation of current, the field effect CRD provides essentially constant current over an almost 100 V variation. It is not surprising then that one useful application for the CRD is in a Zener-CRD combination voltage regulator, as shown in Fig. 22-19. Also shown is the conventional resistor-diode regulator.

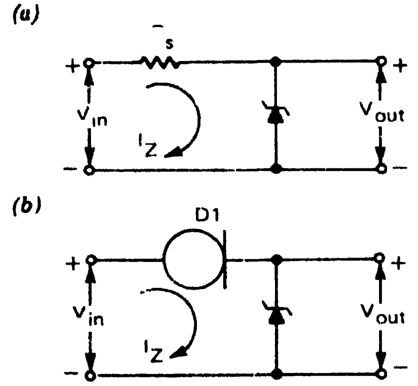


Fig. 22-19. (a) Conventional voltage-regulator circuit. (b) Modified voltage-regulator circuit using a CRD in place of R_s.

There are two distinct advantages of using a CRD instead of a discrete resistor R_s.

(1) If the input voltage V_{in} changes, the change in output voltage, ΔV_{out}, is given by

$$\Delta V_{out} = \Delta V_{in} \frac{Z_{ZT}}{R_s + Z_{ZT}} ,$$

for the conventional circuit, where Z_{ZT} is the dynamic resistance of the Zener diode and is low compared with R_s. The equation for Fig. 22-19(b) may be written as

$$\Delta V_{out} = \Delta V_{in} \frac{Z_{ZT}}{Z_T + Z_{ZT}}$$

and since Z_T may be many megohms compared with only hundreds or thousands of ohms for R_s, the change in V_{out} is much smaller with the CRD.

The above may be interpreted by saying that an increase in V_{in} may be absorbed across D1 *without* an increase in current, since, unlike R_s, the CRD does not obey Ohm's law. This is also the key to the second advantage.

(2) The maximum permissible increase in V_{in} is now limited by the maximum voltage that D1 can absorb (100 V) and not the dissipation capability of the Zener diode. That is, an increase in V_{in} in Fig. 22-19(b) does not cause an increase in Zener current, so the dissipation does not increase either. Only the minimum current demanded by the load determines the maximum dissipation of the Zener diode.

The voltage and current ranges of the CRD may be extended by series and parallel combinations. There are many other interesting applications of the CRD, such as generating sawtooth, triangular and stairstep waveforms, but space does not permit their coverage. However, we shall have occasion to refer to the CRD again in a later chapter, when we utilize its current-regulating ability in a differential amplifier.

22-11. SUMMARY

(1) The MOSFET is a field-effect transistor in which the metal gate is insulated by a very thin oxide layer from the semiconductor channel.

(2) Both depletion and enhancement types of MOSFET are available; in the former, drain current may be controlled by either enhancing or depleting the channel that exists between drain and source, but in the enhancement type drain current flows only when a channel is induced by enhancing carrier concentration.

(3) The major difference between the JFET and the MOSFET is the much larger input resistance in the latter and the handling precautions necessary with those devices that do not have gate protection against static voltages.

(4) A dual-gate MOSFET, as the name implies, provides a means of controlling drain current through the application of two input voltages, one on each gate, important applications of which include AGC, mixer circuits, and demodulation.

(5) Integral gate protection in a dual-gate MOSFET involves two pairs of diodes connected back-to-back between each gate and source. These diodes clip and limit the input voltage to typically ± 10 V, thereby protecting against high static voltages that might puncture the very thin oxide insulating layer.

(6) In addition to having applications in oscillator circuits, the FET, junction or insulated-gate type, can be used in the channel ohmic region as a voltage variable resistor, VVR. In this case, the drain-to-source resistance may be made variable over a very large range by the application of voltage to the gate.

(7) The VVR may be used in an attenuator circuit, an AGC circuit, a chopper-controlled dc amplifier or a solid-state voltmeter, to mention a few applications.

(8) The field-effect diode is essentially a JFET with its gate tied to the source internally to provide a two-terminal current-regulating diode, or CRD.

(9) The constant current of a CRD may be anywhere between 0.2 mA and 5 mA, depending upon the diode selected, and is maintained over a voltage range from approximately 3 V to 100 V.

(10) Although the field effect CRD has many applications involving a constant-current source requirement, it can be used to good advantage in a voltage regulator, with a Zener diode, since the voltage variation at the input will cause essentially no change in output voltage or Zener current.

QUESTIONS

Q22-1. (a) Where is the PN junction located in a P-channel depletion type of MOSFET?
(b) What do the carriers consist of in this type of transistor?
(c) When a positive voltage is applied to the gate of a P-channel depletion MOSFET, is the current flow depleted or enhanced?
(d) Why is it that a MOSFET (depletion type) can work in either the enhancement or depletion mode?
(e) To what *limited* degree can a JFET be operated in the enhancement mode?
(f) Why is a JFET not identified as being of the depletion type?
(g) If a JFET were operated with zero bias (like a depletion-type MOSFET), what is the approximate peak-to-peak voltage it could handle at the input without drawing gate current?

Q22-2. (a) What is the major physical difference between an enhancement-type and a depletion-type MOSFET?
(b) How is this represented in the symbol?
(c) What type of voltage is necessary at the gate of a P-channel MOSFET to cause current flow? What are the induced carriers and where do they come from?

(d) What is meant by the threshold voltage? Is this peculiar to an N-channel enhancement type of MOSFET?

Q22-3. What accounts for the much smaller value of I_{GSS} in a MOSFET compared with a JFET?

Q22-4. Give at least three precautions to observe when handling a MOSFET that does not have any built-in gate protection.

Q22-5. (a) Explain how a dual-gate MOSFET may be thought of as two FETs in series.

(b) Which gate has the signal to be amplified connected to it? What function does the other gate perform?

(c) How does the forward transconductance of Unit 1 vary with gate-to-source voltage in a dual-gate P-channel depletion-type MOSFET? (Include the effects of both gates.)

(d) What effect did the addition of the second gate have on the reverse transfer capacitance from drain to Gate No. 1? Why is this desirable?

Q22-6. Describe how integral gate protection works. What drawback does it have?

Q22-7. Describe the operation of the circuit in Fig. 22-10 as far as maintaining v_o at a near constant level when v_i increases (and V_{AGC} decreases).

Q22-8. (a) At what frequency does the oscillator in Fig. 22-12 operate?

(b) Explain how the transistor's interelectrode capacitances are used to provide feedback.

Q22-9. (a) What is meant by an FET's "channel ohmic region"?

(b) Is the variation of drain resistance with gate voltage linear?

(c) What value of drain-to-source voltage should be used to obtain the effects in the channel ohmic region?

Q22-10. Describe the operation of at least three applications of the FET used as a VVR.

Q22-11. Consider the circuit of the FET voltmeter in Fig. 22-17. Assume that with no voltage applied to the probe the "bridge" is balanced and the meter reads zero (at the center).

(a) When a positive voltage is applied to the probe (switch in Position 2), what voltage is applied to the gate?

(b) For the condition in (a), what will happen to the current through the transistor?

(c) Which way will current flow through the meter, and which way will it deflect for the condition in (a)?

(d) What is the function of the diode? When will it come into play?

(e) Which way will current flow through the meter, and which way will it deflect if a negative voltage is applied to the probe (switch in Position 3)?

(f) Repeat (e) if the switch had been left in Position 2.

Q22-12. (a) What is the principle behind the CRD?

(b) How could you approximate a CRD using an N-channel JFET? Draw a diagram showing the normal application of voltage.

(c) Why does the characteristic curve for the CRD in Fig. 22-18(a) show such a large current in the reverse direction? (See diagram in Q22-12(b).)

Q22-13. Explain how a CRD and a Zener diode combine to provide an excellent voltage-regulator circuit, indicating why it is better than using a discrete series resistor.

PROBLEMS

P22-1. Refer to the circuit in Fig. 22-9.

(a) A MOSFET is being tested, but it is not known which type it is. With the switch open, the light is off, but it comes on after the switch is closed. What type of MOSFET is this?
(b) If the light did not come on after closing the switch in (a), what condition would this indicate?
(c) Specify what polarity of gate voltage should be used to test a P-channel JFET and the indications for (i) good, (ii) open, (iii) shorted.
(d) Repeat (c) for an N-channel enhancement type of MOSFET.

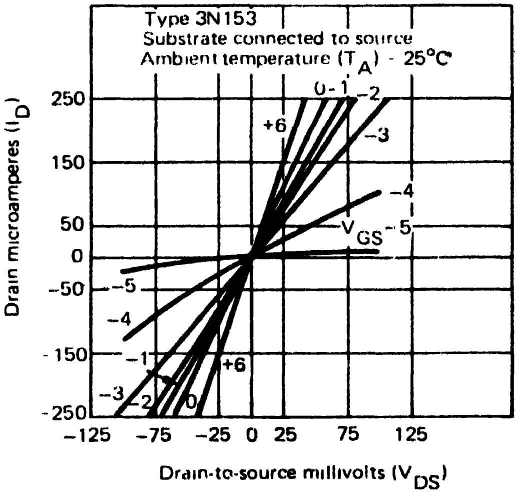
P22-2. A MOSFET amplifier is to be designed with AGC, as in Fig. 22-10, using a 3N200 silicon dual-gate transistor. R_L = 1.2 K ohms. Determine:

(a) The values of bias resistors R_1 and R_2, and the drain supply voltage, V_{DD}, to operate the transistor at a drain-to-source voltage of 15 V at a forward transconductance, g_{fs} = 11,000 μmhos. (Note: There are two values of V_{G1S} that provide this condition, at V_{G2S} = +2 V; carry through the design for both values.) Assume a maximum loading on V_i of 150 K ohms.
(b) The values of bias resistors R_3 and R_4 to provide the conditions in (a) when V_{AGC} = +6 V (maximum loading, 100 K ohms).
(c) The maximum voltage gain of the amplifier, with V_{AGC} = +6 V.
(d) The voltage gain of the amplifier with V_{AGC} = 0 V.
(e) Determine the dB decrease in gain for an AGC variation from +6 V to 0 V. (Which bias point for V_{G1S} is more effective?)
(f) What would be the effect of V_{AGC} becoming -0.5 V?

P22-3. Repeat P22-2 using R_L = 2.2 K ohms and design for a conductance that provides a maximum voltage gain of -28 at V_{DS} = 15 V.

P22-4. Given the low-level drain characteristics of the 3N153 MOS N-channel depletion-type transistor, determine:

(a) The drain-to-source resistance at V_{DS} = 25 mV and V_{GS} = -5 V to 0 V.
(b) Repeat (a) for V_{DS} = 75 mV. Conclusion?
(c) Draw a graph showing how R_{DS} varies with V_{GS}.
(d) If $R_{DS(off)}$ is 10^9 ohms at V_{DS} = 1 V (and V_{GS} = -8 V), find the drain current at this value of gate voltage. Indicate on the above curve in (c) the value of $R_{DS(off)}$.



Typical low-level drain characteristics of a 3N153 MOSFET for P22-4 (Courtesy of RCA Corp.)

P22-5. A voltage-controlled attenuator uses a 3N153 MOS transistor in a circuit like that in Fig. 22-14(a). If R_D = 1 K ohm and V_s = 50 mV p-p, draw a curve showing how the output voltage V_o varies with V_{GS} from 0 to -8 V, using the results of P22-4.

P22-6. Refer to Fig. 22-16. Assume V_{in} is a 1 V p-p 10 Hz sine wave and v_c is a 10 V p-p square wave. Draw the waveform at X for the following frequencies of the square wave (assume the square wave initially goes negative at the instant the sine wave is going positive):

(a) 20 Hz, (b) 10 Hz, (c) 40 Hz, (d) 5 Hz.

P22-7. Refer to Fig. 22-15. Assume Q1 has an $h_{ie} = 2$ K ohms and $h_{fe} = 100$. Assume that when $v_s = 1.2$ V p-p R_e is adjusted to 4 K ohms and the filter-clamper circuit at the output produces -2.8 V at the gate of Q2, making its effective drain-to-source resistance 4 K ohms.

(a) Determine the voltage gain of the amplifier and the output voltage.
(b) If v_s increases to 1.5 V p-p, what would be the new output voltage if there were no AGC action (resistance of Q2 still 4 K ohms)?
(c) If the increase of v_s in (b) increased the negative voltage at the gate of Q2 to -3 V, making its effective resistance 6 K ohms, calculate the new voltage gain of the amplifier and the new output voltage. Compare with (b).
(d) Could the AGC action ever be so "strong" that it could cause a reduction in v_o for an increase in v_s? Explain.

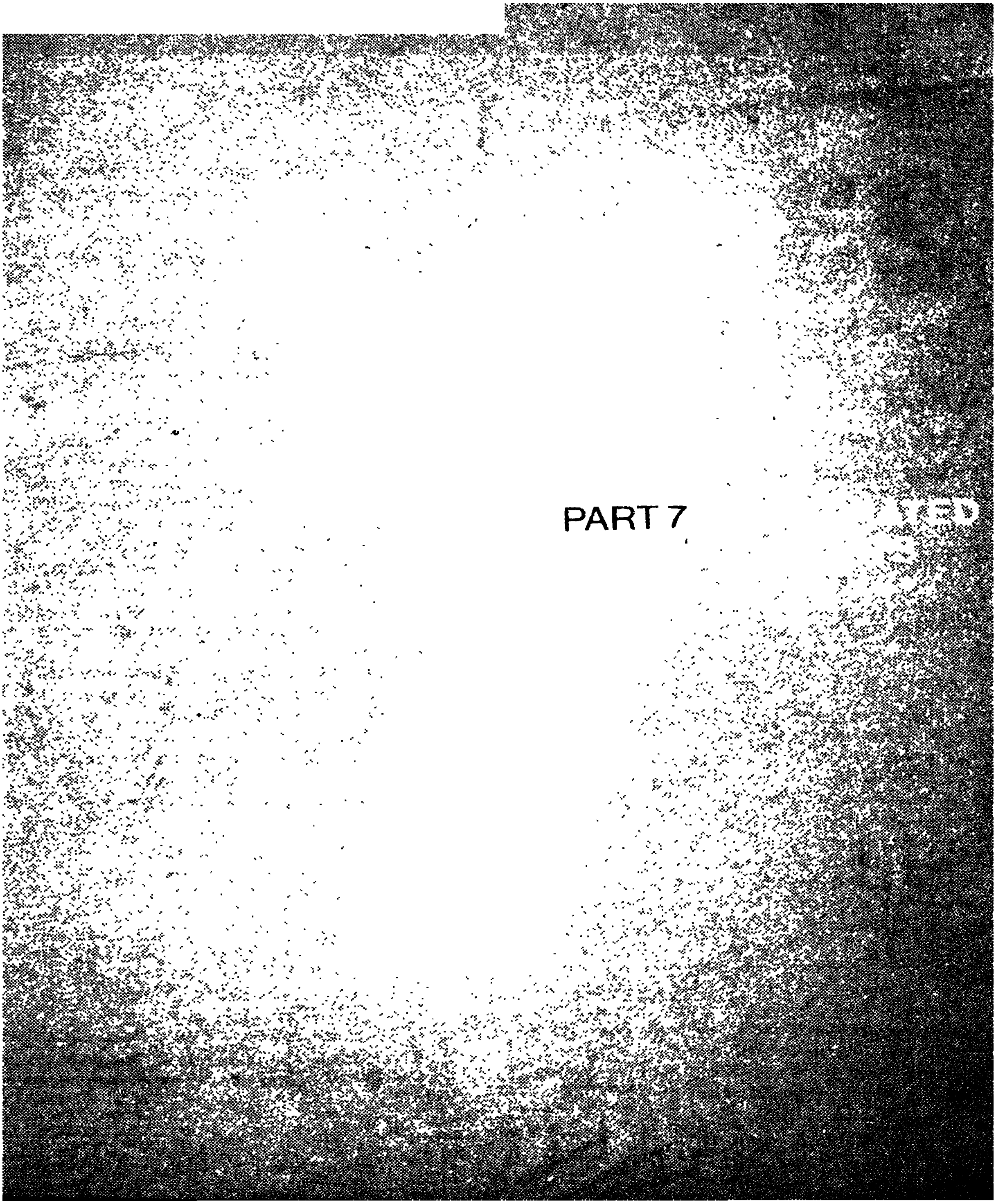
P22-8. Refer to the voltage regulator circuits in Fig. 22-19. In the conventional circuit, an 18 V Zener diode is used with a 30 V input to provide voltage control to a load whose current varies from 0 to 2 mA.

(a) Using a Zener current of 2 mA with an average load current of 1 mA, determine the necessary size of R_s.
(b) If the Zener diode has a dynamic impedance of $Z_{ZT} = 40$ ohms, determine the change in output voltage if V_{in} changes from 30 V to (i) 32 V, (ii) 60 V.
(c) Determine the maximum dissipation of the Zener diode when V_{in} increases to 60 V.

Consider now that R_s is replaced by a 1N5309 field-effect diode, or CRD, that has the characteristics shown in Fig. 22-18. $Z_T = 300$ K ohms.

(d) Repeat part (b) and compare.
(e) Repeat part (c) and compare.
(f) What is the maximum voltage at the input that the circuit with the CRD can withstand? What limits the voltage in the conventional circuit?
(g) What is the minimum permissible voltage at the input that will maintain voltage regulation?

P22-9. Repeat P22-8 using a load current variation from 0 to 4 mA; a Zener current of 4 mA with an average load current of 2 mA (for part a); two 1N5309 CRDs are used in parallel to provide the necessary current capability (for parts d through g).



PART 7

23

LINEAR INTEGRATED CIRCUITS: OPERATIONAL AMPLIFIERS

23-0 INTRODUCTION

Linear integrated circuits are found in four main areas:

(1) Power amplifiers (audio applications with 1 watt or more of output power)

(2) High frequency amplifiers (rf, video and wideband, dc to 75 MHz)

(3) Differential operational amplifiers (high input resistance and open loop gain)

(4) Voltage regulators (for example, up to 0.6 A at 40 V; low output impedance of 20 m ohms)

In this chapter, we will concentrate on the differential operational amplifier, since it exemplifies the type of circuit that is used in all of the above areas. This circuit is the difference or differential amplifier, which is particularly suited to integrated-circuit technology. A number of applications of the differential operational amplifier will be given, including that of an analog computer circuit, a voltage comparator, temperature measuring circuit, and a Wien-bridge oscillator. A voltage regulator that utilizes an operational amplifier is also examined, as is an audio power amplifier.

Since the fabrication of integrated circuits will not be covered until the next chapter, we will consider the IC here mainly as a "black-box" device, only occasionally referring to its schematic structure. However, basically, a small monolithic integrated circuit consists of a single chip of silicon, typically 40 by 40 mils in area, in which both active and passive components have been diffused and then interconnected. Hybrid circuits also exist that are a mixture of diffused and thin film devices. ICs may be packaged in circular metal "cans" with up to 12 leads, in plastic "in-line" packages with up to 16 leads, or in flat packages with over 32 leads.

23-1. THE OPERATIONAL AMPLIFIER

Originally used to perform mathematical functions such as addition, sign changing, integration, differentiation, etc., in analog computers, the operational amplifier is now found in many other applications, notably in control systems. Basically, it consists of a very high gain dc amplifier with high input impedance and low output impedance. The early operational amplifiers had only one input, and the output was always inverted with respect to the input. However, general purpose operational amplifiers now available are usually of the differential type, where both inverting and noninverting inputs are available. See Fig. 23-1.

It is understood that all voltages are with respect to ground, so the ground line is not usually shown. In Fig. 23-1(a), $V_o = -AV_s$; in Fig. 23-1(b), $V_o = AV_2 - AV_1 = A(V_2 - V_1)$, where V_2 and V_1 are measured with respect to ground. Thus, A is the differential voltage gain and will later be given the notation A_d, since it

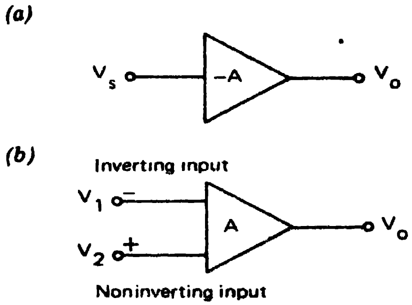


Fig. 23-1. (a) Single-input operational amplifier with negative gain. (b) Dual-input differential operational amplifier.

amplifies the difference in input voltages. Of course, if either V_1 or V_2 is at ground level, as in many operational applications, $V_o = AV_2$ or $-AV_1$, respectively.

The amplifier's gain, A, is also referred to as the open loop dc gain and may be anywhere from 60 dB (1,000) to 140 dB (10 million), so they are used in circuits with considerable negative feedback.

Let us now consider the characteristics and applications of the inverting differential operational amplifier.

23-2. THE INVERTING DIFFERENTIAL OPERATIONAL AMPLIFIER

Gain

Let us determine the overall voltage gain of the amplifier shown in Fig. 23-2. V_o/V_1 is known as the closed loop voltage gain.

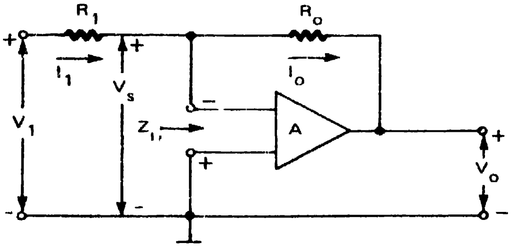


Fig. 23-2. Operational amplifier with negative feedback, to be analyzed for overall voltage gain V_o/V_1.

Note that V_1 may be a dc voltage or an ac signal within the bandwidth of the amplifier, which may be as high as 1 MHz.

The differential input impedance between the two input terminals, Z_i, is ideally infinite but is considered large enough compared with R_o so that $I_1 = I_o$. That is, negligible current flows into the operational amplifier.

But

$$I_1 = \frac{V_1 - V_s}{R_1}, \text{ and } I_o = \frac{V_s - V_o}{R_o}$$

$$\begin{matrix} V_1 - V_s \\ R_1 \end{matrix} \quad \frac{V_s - V_o}{R_o}$$

$$\therefore \frac{V_1}{R_1} \quad \frac{V_s}{R_1} \quad R_o \,^{-}\, R_c$$

$$\frac{V_o}{R_o} = \frac{V_s}{R_o} + \frac{V_r}{R} \quad R_1$$

$$= V_{-}\left(\frac{1}{R_o} + \frac{1}{R_1}\right) - \frac{V_1}{R_1}$$

But $V_o = -AV_s$

$$\therefore V_s = \frac{-V_o}{A},$$

$$\therefore \frac{V_o}{R_o} = \frac{-V_o}{A}\left(\frac{1}{R_o} + \frac{1}{R_1}\right) - \frac{V_1}{R_1}. \qquad (23\text{-}1)$$

While this can be solved for V_o/V_1, let us make the following observation:

V_o is typically 10 V maximum, A is at least 10,000, and R_o and R_1 are often in the megohm range but not less than 10 K ohms.

Thus a typical value for

$$\frac{V_o}{A}\left(\frac{1}{R_o} + \frac{1}{R_1}\right) \simeq \frac{-10}{10^4}\left(\frac{1}{10^4} + \frac{1}{10^4}\right)$$

$$-2 \times 10^{-7},$$

while a typical value for

$$\frac{-V_1}{R} \simeq \frac{-10}{10^4} \simeq -1 \times 10^{-3}.$$

Thus, the first term is negligible compared with the second (a factor of 5,000 times as small), so to a good degree of accuracy we have

$$R_{\_} \quad \frac{-V_1}{R_1}, \text{ and } V_o \quad \frac{-R_c}{R_1} \quad I_1, \text{ or } \frac{V_o}{V_1} = \frac{-R_o}{R_1}. \tag{23-2}$$

If R_o = 100 K ohms and R_1 = 10 K ohms, $V_o = -10\ V_1$, and an amplifier with phase inversion results, having a closed loop gain of 10 that is dependent only upon the feedback components. The question may be asked, will not the same gain result if R_o = 1 M ohm and R_1 = 100 K ohms? This will depend upon the input impedance to the operational amplifier, Z_i. Many discrete FET operational amplifiers have $Z_i = 10^{12}$ ohms, so R_o could be many megohms without disturbing our assumption that $I_1 = I_o$. But in many IC amplifiers Z_i seldom exceeds 1 M ohm and is typically less than 0.5 M ohm. (IC amplifiers utilizing FET inputs are an exception, with Z_i typically 10^{12} ohms.) As a result, R_o (and R_1) must be selected in accordance with Z_i if the above equation for gain is expected to be valid.

Input Impedance

In the derivation above, we used

$$V_s = \frac{-V_o}{A}.$$

For a value of A very large and near infinity, V_s becomes essentially zero, and the negative input terminal in the inverting amplifier above is often referred to as a "virtual ground." It is also called the summing point, as will be seen in the next application, and accounts for using V_s for the notation. As a result of this virtual ground, it should be plain that the overall input impedance to the amplifier, V_1/I_1, is R_1.

Output Impedance

The open-loop output impedance of the operational amplifier may be anywhere from 1 ohm in high current amplifiers to 10 K ohms in FET amplifiers. Typical IC amplifiers have a value of 5 K ohms. The closed-loop output impedance (with feedback) is less than the open-loop output impedance (no feedback) according to the following equation:

$$Z_{o_{CL}} \quad \frac{Z_{o_{OL}}}{\text{Loop Gain}}. \tag{23-3}$$

We need now a definition of loop gain. This is defined as follows:

Loop Gain (dB) ≃ Open Loop Gain (dB) - Closed Loop Gain (dB). (23-4)

As an example, consider an operational amplifier whose output impedance is 5 K ohms and its open loop gain is 80 dB. Feedback components R_1 = 10 K ohms and R_o = 100 K ohms are connected to give a closed loop gain of 10. This is a dB gain of 20 log 10 = 20 dB.

The loop gain is then 80 - 20 = 60 dB, which is a ratio of 1000.

$$Z_{o_{CL}} = \frac{Z_{o_{OL}}}{\text{Loop Gain}} = \frac{5\text{ K ohms}}{1000.} = 5\text{ ohms.}$$

(The input impedance would be 10 K.)

Summing Amplifier

Now consider the application of a number of input voltages to the inverting input terminal, as shown in Fig. 23-3. In a similar manner to the above derivation, we can show that

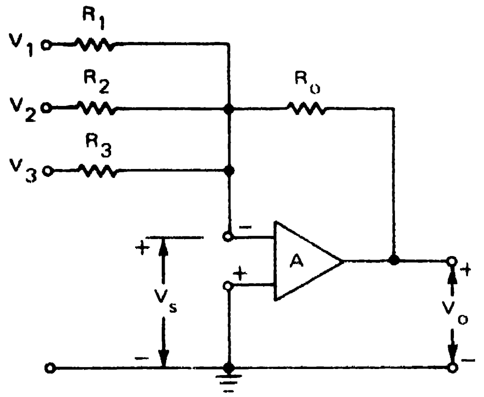


Fig. 23-3. Showing how an operational amplifier may be used to sum a number of input voltages.

$$V_o = -\left(\frac{R_o}{R_1}V_1 + \frac{R_o}{R_2}V_2 + \frac{R_o}{R_3}V_3\right). \qquad (23\text{-}5)$$

If we make $R_1 = R_2 = R_3 = R_o$, then $V_o = -(V_1 + V_2 + V_3)$, and the virtual ground is truly a summing point of the input currents. This connection is very useful in analog computers and may be used to solve linear equations.

23-3. ANALOG COMPUTER SOLUTION OF LINEAR EQUATIONS

Consider solving the following set of equations using operational amplifiers:

$$3X + 6Y - 10 = 0 \qquad (1)$$

$$2X - Y - 8 = 0. \qquad (2)$$

Solving (1) for X, $3X = 10 - 6Y$

$$\therefore X = \frac{10}{3} - 2Y = -\left(\frac{-10}{3} + 2Y\right). \qquad (3)$$

Solving (2) for Y,

$$Y = 2X - 8 = -(8 - 2X). \qquad (4)$$

(1) Assume a voltage representing Y is available, so that we may develop a voltage representing X as shown in Fig. 23-4(a). That is,

$$X = -\left(\frac{R_o}{R_1}(-1) + \frac{R_o}{R_2}(Y)\right) = -\left(\frac{-10\text{ K}}{3\text{ K}} + \frac{10\text{ K}}{5\text{ K}}Y\right)$$

$$= -\left(\frac{-10}{3} + 2Y\right)$$

$$\frac{10}{3} - 2Y.$$

(2) To obtain Y we need a voltage representing -X. Assume this is available (by negating or inverting X), so that Y is developed as shown in Fig. 23-4(b). That is,

$$Y = -\left(\frac{R_o}{R_1}(+4) + \frac{R_o}{R_2}(-X)\right)$$

$$= -\left(\frac{2\text{ K}}{1\text{ K}}(+4) + \frac{2\text{ K}}{1\text{ K}}(-X)\right)$$

$$= -(8 - 2X) = -8 + 2X.$$

The completed analog computer circuit is shown in Fig. 23-4(c).

To reduce loading on the previous stage (which may introduce errors), all resistors could be made five times as large.

The solution to the equations would be indicated by the readings of two dc voltmeters connected at X and Y. A mathematical solution would yield X = 3-13/15 and Y = -4/15.

The accuracy of the voltage representations would depend upon the resistance ratios, but values of +3.85 volts at X and -0.25 volts at Y would be reasonable. In a commercial analog computer, high resolution potentiometers are available to give the desired accuracy of the resistors and input voltages.

The Integrating Amplifier

More complex equations may be solved if an integrating amplifier is used as shown in Fig. 23-5.

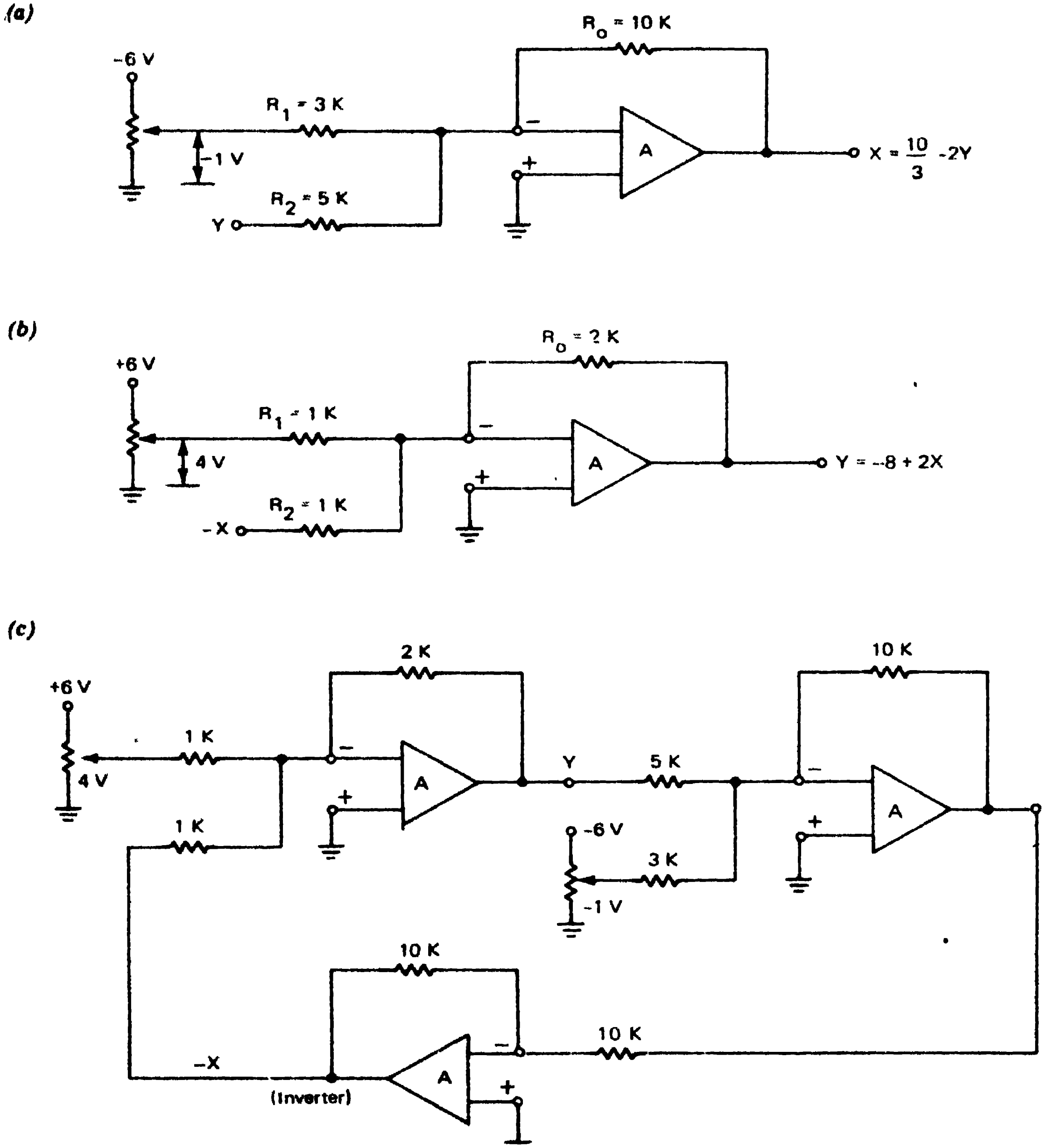

(a)
-6 V
R1 = 3 K
-1 V
R2 = 5 K
Y
Ro = 10 K
A
X = 10/3 -2Y
(b)
+6 V
R1 = 1 K
4 V
R2 = 1 K
-X
Ro = 2 K
A
Y = -8 + 2X
(c)
+6 V
1 K
4 V
1 K
2 K
A
Y
5 K
-6 V
3 K
-1 V
10 K
A
10 K
-X
A
(Inverter)
10 K

Fig. 23-4. (a) Development of X on the assumption Y is available. (b) Development of Y on the assumption -X is available. (c) The complete analog computer circuit in which -X is obtained by inversion of X. Solutions would be read by dc voltmeters at X and Y.

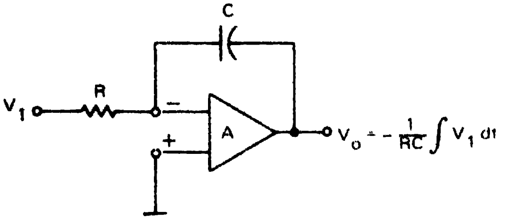
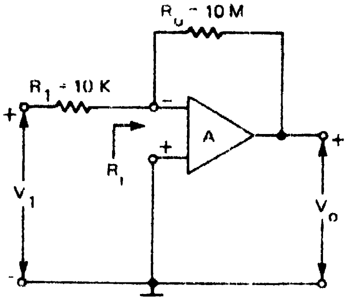


Fig. 23-5. Circuit of an operational amplifier used as an integrator.

This allows the solution of differential equations from the calculus. These equations are beyond the scope of this text, but the integrating amplifier permits a dynamic solution to be displayed on an oscilloscope and allows the simulation of many mechanical problems by using an equivalent electrical circuit or "analog." Contrary to popular belief, the digital computer has not made the analog computer obsolete. It is still a very powerful tool in analyzing both linear and nonlinear control systems, usually at much lower cost than with a digital computer.

23-4. INCREASING INPUT IMPEDANCE IN AN INVERTING AMPLIFIER

Consider trying to obtain an amplifier with a gain of 1000 and an input resistance of at least 10 K ohms from an operational amplifier whose differential input impedance is 500 K ohms and A is 10,000. This means R_1 would be 10 K ohms and R_o would have to be 10 M ohms in the circuit of Fig. 23-6. However, since R_o is now greater than the input impedance of the operational amplifier, the gain equation of R_o/R_1 is no longer valid. It now is necessary to use the exact equation for gain:



If R_i >>10 M ohms

$$\frac{V_o}{V_1} = -\frac{R_o}{R_1} = -\frac{10M}{10K} = -1000$$

Input resistance = R_1 = 10 K

Fig. 23-6. Circuit of an operational amplifier that gives a gain of 1000 with an input resistance of 10 K, if R_i >> 10 M ohms.

$$\frac{V_o}{V_1} = \frac{\dfrac{-R_o}{R_1}}{1 + \dfrac{1}{A'}\left(\dfrac{R_o + R_1}{R_1}\right)}, \qquad (23\text{-}6)$$

where A' = effective open loop gain

$$= \frac{A}{1 + \dfrac{R_oR_1}{R_i\,(R_o + R_1)}}$$

and $R_i = Z_i$ = differential input impedance.

In this case, $\frac{V_o}{V_1}$ would be -900, not -1000.

In order to get the required gain of 1000, it would be necessary to decrease R_1 below 10 K ohms.

One way of increasing the input impedance is to use an extra voltage divider on the output, as indicated in Fig. 23-7.

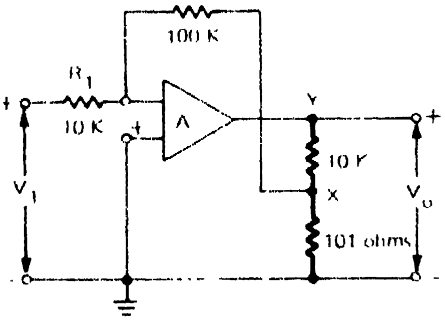


Fig 23-7 Showing how a voltage divider at the output can increase the gain by reducing negative feedback.

The voltage developed at X is now 0.01 V_o. This reduces the feedback to one hundredth of what it would be if connection had been made to Y directly. Thus, a gain of 1000 is obtained, since R_o is now less than Z_{in}. In fact, it would now be possible to make R_1 = 100 K ohms and R_o = 1 M ohm, and still get very close to 1000 in gain, but the input impedance would now be 100 K ohms.

23-5. THE NONINVERTING DIFFERENTIAL OPERATIONAL AMPLIFIER

An excellent way to increase input impedance is to use the noninverting connection shown in Fig. 23-8.

Gain

Here, voltage is applied to the noninverting input, so no phase inversion results at the output. Some of the output is fed back to the inverting input, providing effectively some voltage V_1 at the inverting terminal of the same polarity as at the noninverting input.

Again we assume that $I_o = I_1$, from which

$$V_o = \left(\frac{R_1 + R_0}{R_1}\right) V_2, \qquad (23\text{-}7)$$

and considerable gain can be obtained without phase reversal. The above equation assumes an amplifier gain A of infinity.

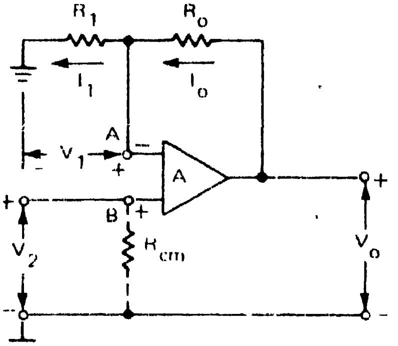


Fig. 23-8. Circuit of a noninverting differential operational amplifier to increase input impedance.

If a more exact value of voltage gain is needed, the following can be used which takes into account a finite value for A:

$$\frac{V_o}{V_2} = \frac{\dfrac{R_1 + R_o}{R}}{1 + \dfrac{1}{A} \times \dfrac{R_1 + R_o}{R_1}} \qquad (23\text{-}8)$$

If R_1 = 10 K ohms, R_o = 100 K ohms, and A = 10,000, the first equation gives a gain of 11, the second equation a gain of 10.99.

Input Impedance

It can be seen, for the above differential amplifier, that

$$V_o = A(V_2 - V_1), \text{ or } V_2 - V_1 = \frac{V_o}{A}. \qquad (23\text{-}9)$$

For large values of A, V_o/A becomes nearly zero, and $V_2 = V_1$. In other words, the operation of the differential amplifier is such as to maintain approximately the same voltage at the two input terminals to the amplifier. The amplifier is said to be operating with some *common mode* voltage at its input terminals, because some voltage is common to both terminals. Thus, there is effectively no voltage drop across the input terminals A and B, because the voltage at A follows that at B, and effectively an open circuit (like a bootstrapped circuit) results between these two points. Hence, the input resistance as far as V_2 is concerned is determined primarily by the common mode input impedance, R_{cm}, and is essentially the leakage resistance between one input and ground. It is much higher than R_i. In a typical IC operational amplifier R_{cm} is typically 200 M ohms compared with R_i = 300 K ohms.

The actual input impedance consists of R_{cm} in parallel with (R_i × loop gain). (Recall that Loop Gain = Open Loop Gain ÷ Closed Loop Gain for voltage ratios.) Thus, if loop gain is small (large closed loop gain), this can cause appreciable extra loading in some amplifiers. Nevertheless, the noninverting amplifier will always provide a much higher input impedance than the inverting amplifier, usually in excess of 1 M ohm and up to thousands of megohms.

It should be mentioned, where input impedance is critical, that *instrumentation amplifiers* are available. These consist of selected IC operational amplifier pairs, matched for equal characteristics and connected for very high differential and common-mode input impedances.

Output Impedance

The output impedance is reduced in the same manner as for the inverting amplifier.

All of the circuits involving the inverting amplifier always had the noninverting input tied to ground. Thus, the amplifier was used as a single-input device – as a true operational amplifier.

The high input impedance in the circuit of Fig. 23-8 came about through the use of both inputs; neither was directly grounded. Let us now examine the relationship between the output voltage and the two input voltages to a differential amplifier.

23-6. THE DIFFERENTIAL AMPLIFIER

Sometimes it is necessary to amplify the difference between two voltage levels and reject the voltage common to both, as shown in the circuit of Fig. 23-9.

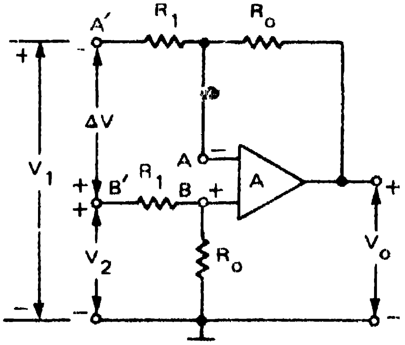


Fig. 23-9. Circuit of a differential amplifier to provide gain for $V_2 - V_1$.

It may be shown that, ideally,

$$V_o = \frac{R_o}{R_1}\Delta V, \text{ where } \Delta V = V_2 - V_1. \qquad (23\text{-}10)$$

The overall amplifier operates in such a manner as to maintain near zero volts between A and B, with the result that the differential input impedance for ΔV between A′ and B′ is $2R_1$. This loading must be considered in some of the following applications.

Thus, if R_o = 100 K ohms, R_1 = 10 K ohms, V_2 = +3.1 volts, and V_1 = +3.0 volts, then V_o = 10(3.1 -3.0) = 1.0 V, and the differential input impedance is 20 K ohms.

In practice, the amplifier operating in the differential mode with significant common mode voltage at the input terminals will provide an output that is dependent not only on the difference of the two input signals but also upon the common mode input voltage. This is due to the difference in gain for each signal, as explained in the next section.

23-7. A GENERAL PURPOSE IC OPERATIONAL AMPLIFIER

Refer to the circuit of a differential operational amplifier shown in Fig. 23-10.

This is a monolithic circuit featuring Darlington-connected NPN transistors for high input impedance and two differential gain stages with a single-ended emitter-follower output. This is a popular circuit for ICs, since it is possible to very closely match the characteristics of transistors Q1-Q2 and Q3-Q4 and resistors R_1-R_2. Their close spacing on the chip reduces errors arising from thermal drift, since the two components of each pair are affected equally.

Consider A to be grounded and a small positive voltage to be applied at B. Point Y will become less positive, Q4 will turn on more, and Z will become less negative. Q5 will conduct less, so that the output will rise to some positive value. This will provide some gain, A_1. Similarly, if B is grounded and the same positive voltage is applied to A, X will become less positive; Q3 will conduct more, making K less positive; Q4 will conduct less, making Z more negative. The change in voltage at the collector of Q4 is transferred to the output as a similar negative voltage change, which provides some gain A_2. If the two voltages are applied simultaneously at A and B and if A_1 = A_2, then the output will be zero. (This assumes that the offset voltage appearing at the output when the two inputs are both zero has been nulled by adjusting a suitable

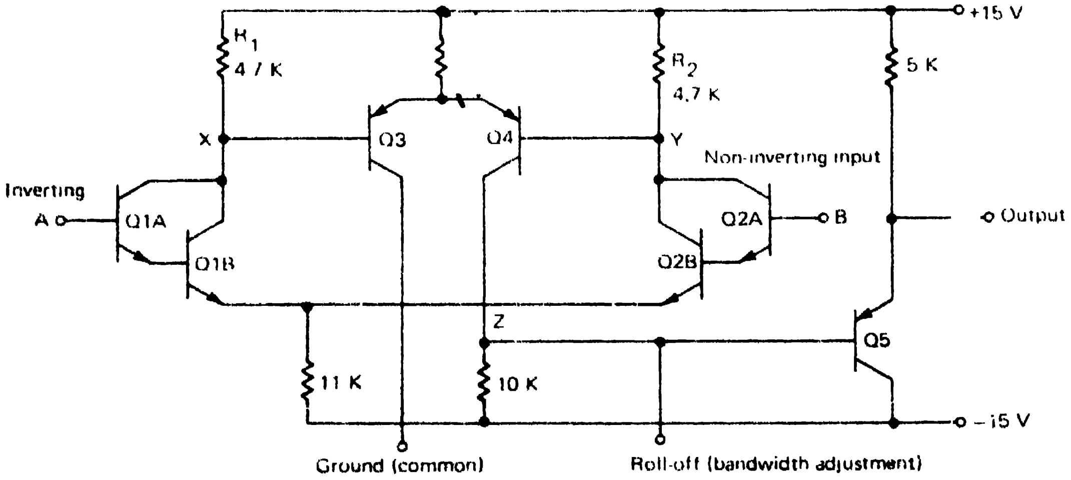


Fig. 23-10. Circuit of a monolithic IC differential operational amplifier.

balance resistor as directed by the manufacturer. This is necessary to overcome any slight unequal base-emitter bias voltages that may exist in the input transistors.)

However, A_1 will not always be equal to A_2 due to slight differences in components, and an output will result even though the two inputs are identical. The ratio of the output to the common mode input (when the two inputs are equal) is called the common mode gain A_c and would be zero in an ideal differential operational amplifier. When the two inputs are not equal, the common mode voltage is given by

$$V_1 + V_2 \tag{23-11}$$

The desired gain of the amplifier is the differential gain, A_d, (or open loop gain) and is the ratio of the output to the differential input ΔV, with one of the inputs grounded.

23-8. COMMON-MODE REJECTION RATIO

To express how successful an amplifier is in providing gain for the differential signal and rejecting the common mode signal, a factor called the common-mode rejection ratio is defined:

$$\text{CMRR} \quad \frac{A_d}{A_c} = \frac{\text{Differential Gain}}{\text{Common-mode Gain}}. \tag{23-12}$$

Alternatively, the common-mode rejection may be expressed in decibels as:

$$\begin{aligned}\text{CMR} &= 20 \log \text{CMRR} \\ &= 20 \log A_d - 20 \log A_c. \end{aligned} \tag{23-13}$$

For example, in the above IC amplifier of Fig. 23-10 the CMR is typically 55 dB, and A_d is typically 1200. What are A_c and CMRR?

$$20 \log A_d = 20 \log 1200 = 20 \times 3.08 = 61.6\,\text{dB}$$
$$20 \log A_c = 20 \log A_d - \text{CMR}$$
$$\therefore 20 \log A_c = 61.6 - 55 = 6.6\,\text{dB}$$

$\therefore A_c$ = antilog 0.33 2.14 and the

$$\text{CMRR} = \frac{A_d}{A_c} = \frac{1200}{2.14} = 560.$$

As an example, if 1.5 mV is applied to B and 1.3 mV to A in the circuit of Fig. 23-9 (at the amplifier terminals), the output due to differential gain will be $(V_2 - V_1)A_d = (1.5 - 1.3)\text{ mV} \times 1200 = 0.24\text{ V}$.
To this must be added the output due to common mode gain, $V_{cm} A_c$:

$$\left(\frac{V_1 + V_2}{2}\right) \times 2.14 = \frac{1.5 + 1.3}{2}\text{ mV} \times 2.14$$
$$= 3\text{ mV}.$$

Therefore, the total output will be 0.243 volt. However, if the inputs were 5.0002 V and 5.0000 V then the differential output would be (5.0002 - 5.0000) 1200 = 0.24 volt, and that due to the common mode gain 5.0001 × 2.14 = 10.7 V.

This has introduced a large error and may cause the amplifier to saturate. Fortunately, when the amplifier is used with feedback, the overall common-mode gain is reduced, as shown by the following. Refer to Fig. 23-11.

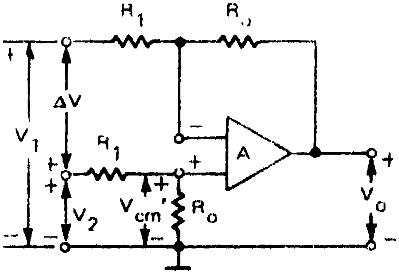


Fig. 23-11. Circuit of a differential operational amplifier showing the common mode input voltage V_{cm}'.

The common-mode signal for the overall amplifier is

$$V \;=\; \frac{V_1 + V_2}{}$$

The common-mode signal appearing at the input-terminals of the differential amplifier is

$$V'_{cm} = V_2 \frac{R_o}{R_o + R_1}. \tag{23-14}$$

The total output voltage is then given by

$$V_o = \frac{R_o}{R_1}\Delta V + \underbrace{\frac{R_o}{R_1} V'_{cm} \frac{1}{CMRR}}_{\text{error term}}, \tag{23-15}$$

where: $\Delta V = V_2 - V_1$.

Obviously, a large CMRR will reduce the error term. As an example, let

$R_o = 100$ K ohms, $R_1 = 10$ K ohms, $V_2 = 5.5$ V, and $V_1 = 5.1$ V.

$$\text{Then } V_o = \frac{100}{10}(5.5 - 5.1)$$

$$+ \frac{100}{10} 5.5 \left(\frac{100}{100 + 10}\right)\frac{1}{560}$$

$$= 4 + \frac{50}{560} = 4 + 0.09 = 4.09 \text{ volts.}$$

The desired output in this case is 4 V, and the error due to the common mode voltage is 0.09 V. In order to increase the CMRR to reduce this error, we need an amplifier with a larger differential gain A_d and a lower common-mode gain A_c. Let us briefly look at the circuit factors that determine the value of the CMRR.

23-9. EMITTER-COUPLED DIFFERENTIAL AMPLIFIER

Refer to Fig. 23-12(a), where transistors Q1 and Q2 have their emitters coupled to a common resistor R_e. Consider $V_{o2} = -I_{c2}R_{L2} = V_o$ to be the output voltage with respect to ground. If the amplifier is followed by additional stages of gain, both V_{o1} and V_{o2} would be used to drive the succeeding stages, as will be shown later.

Common-Mode Voltage Gain, A_c

Let $V_1 = V_2 = V_s$. Then $A_c = \dfrac{V_o}{V_s}$.

Due to symmetry, each input at the base sees a common emitter circuit with an unbypassed emitter resistor of $2R_e$. (The emitter resistor is effectively doubled, since it carries the emitter current for both transistors.) From Chapter 14, we have:

$$R_i = h_{ie} + (1 + h_{fe})2R_e$$

$$A_i = -h_{fe}$$

$$\therefore A_v = \frac{A_i R_L}{R_i}$$

$$= \frac{-h_{fe}R_L}{h_{ie} + (1 + h_{fe})2R_e} \approx \frac{-R_L}{2R_e} = A_c, \tag{23-16}$$

since usually $(1 + h_{fe})2R_e >> h_{ie}$, and $1 + h_{fe} \approx h_{fe}$.

Differential-Mode Voltage Gain, A_d

Let $-V_1 = V_2 = V_s/2$. That is, the input voltage between B and A, $V_{BA} = V_2 - V_1 = V_s/2 - (-V_s/2) = V_s$.

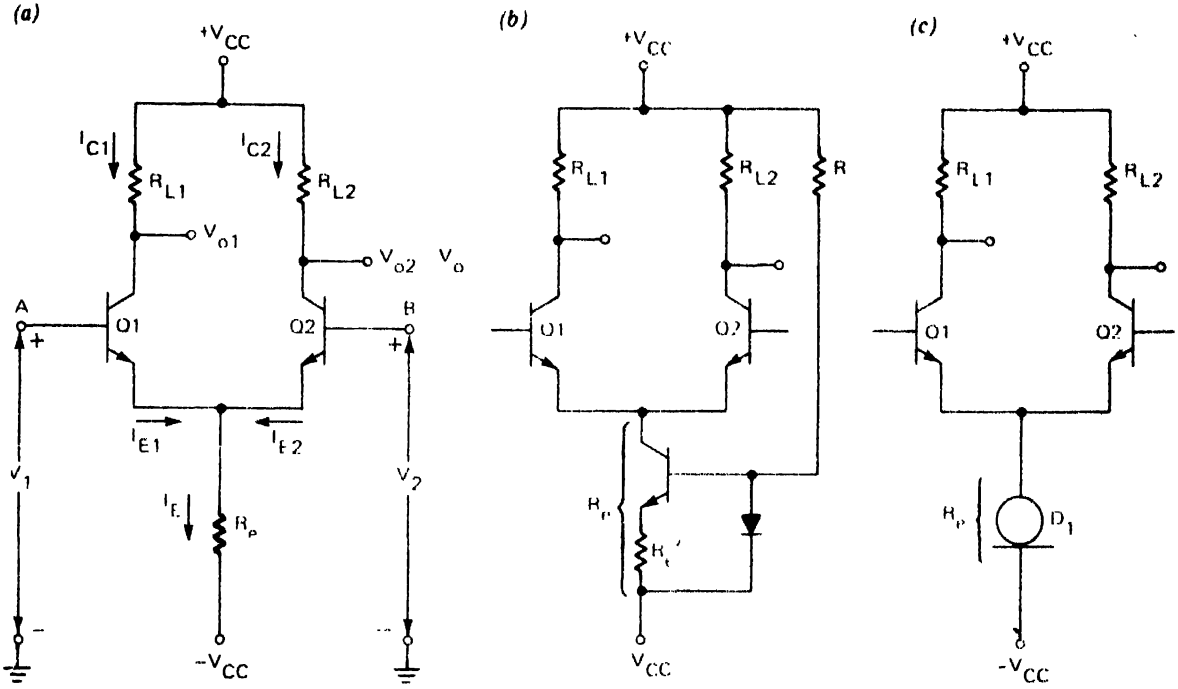

(a)
+V_CC
I_C1
R_L1
I_C2
R_L2
V_o1
V_o2
V_o
A
+
Q1
Q2
B
+
I_E1
I_E2
V_1
V_2
I_E
R_e
−
−
−V_CC
(b)
+V_CC
R_L1
R_L2
R
Q1
Q2
R_e
R_e'
V_CC
(c)
+V_CC
R_L1
R_L2
Q1
Q2
R_e
D_1
−V_CC

Fig. 23-12. (a) Emitter-coupled differential amplifier with emitter resistor R_e. (b) R_e provided by a constant-current transistor. (c) Current-regulating diode (CRD) used to provide a high effective value for R_e.

Then $A_d = \frac{v}{V_2 - V_1} \quad \frac{V_o}{V_-}$.

The total input resistance seen between the bases by the input signal is, due to symmetry, given by: $R_i = 2h_{ie}$; $A_i = -h_{fe}$.

$$\therefore A_v = \frac{A_i R_L}{R_i} \quad \frac{-h_{fe} R_L}{2h_{ie}} = A_d \qquad (23\text{-}17)$$

Common-Mode Rejection Ratio, CMRR

$$\text{CMRR} = \frac{A_d}{A_c} \quad \frac{\frac{-h_{fe} R_L}{2h_{ie}}}{2R_e} \quad \frac{h_{fe} R_e}{h_{ie}} \qquad (23\text{-}18)$$

If in Fig. 23-10 with R_e = 11 K ohms, h_{fe} = 50 and h_{ie} = 1 K ohm, then

$$\text{CMRR} = \frac{h_{fe}R_e}{h_{ie}} \doteq \frac{50 \times 11 \times 10^3}{1 \times 10^3} = 550,$$

which is close to the value of 560 determined in Section 23-8.

Constant Current Operation

Equation 23-18 points up the desirability of as large a value of R_e as is practical to provide a high degree of common-mode rejection. However, the voltage drop across R_e due to the total emitter current, I_E, limits the size of R_e to avoid impractically high values of $-V_{CC}$. There are two methods of increasing the effective value of R_e.

(1) A transistor in the CE configuration with an unbypassed emitter resistor, R_e', has a very high output resistance between collector and emitter, approaching the value of $1/h_{ob}$. This modification is shown in Fig. 23-12 (b). The diode is often added to stabilize the bias of the transistor by compensating for the variations in V_{BE} with temperature change. This is especially effective in integrated circuits, since the close proximity of the two devices will cause both junctions to be equally affected. Amplifiers utilizing a transistor for R_e, such as that in Fig. 23-13. have a CMR $\simeq$ 90 dB, or a CMRR $\simeq$ 31,600.

(2) It should be evident from the foregoing that a large effective value for R_e is an attempt to produce a constant current source for I_E. A device that admirably fills this requirement is the field-effect diode or CRD, the current regulating diode, discussed in Section 22-10. See Fig. 23-12(c). It can present an effective resistance of up to 25 M ohms with as little as 1 V drop across it.

It should be noted that a constant-current source in the emitter circuit means that collector current in one transistor can increase only at the expense of collector current in the other transistor. Therefore, the collector voltage of Q2 in Fig. 23-12, for example, will increase, while the voltage at the collector of Q1 will decrease by an equal amount, due to some differential voltage applied between B and A. This is precisely what is required. The differential gain, A_d, if the output is taken between the collectors of Q1 and Q2, will be double the value obtained between one collector and ground.

If one of the inputs to the differential amplifier is grounded and a sinusoidal signal is applied to the other, a phase-splitting function is achieved, where the output voltages at the two collectors are 180° out of phase with each other with respect to ground.

23-10. HIGH-PERFORMANCE OPERATIONAL AMPLIFIER

The schematic of an IC high-performance operational amplifier is shown in Fig. 23-13.

The emitter-coupled differential-input amplifier and its transistorized current source should be evident. The outputs from the collectors drive a balanced pair of modified Darlington pairs. Further amplification is achieved by using two stages of complementary symmetry amplifiers along with a common emitter. In this way, a single-ended output is obtained. This amplifier has a typical open loop gain A_d = 45,000, CMR = 90 dB, R_{in} = 400 K ohms, R_{out} = 150 ohms; it can provide a ±13 V peak output with a 2 K ohm load when supplied by ±15 V. It can accept common-mode input voltages of ±10 V. Its rise-time capability t_r = 0.3 μsec with a 20 mV input. Terminals labeled as input and output lag provide a means of restricting the amplifier's bandwidth by external compensating R-C networks specified by the manufacturer. This adds amplifier stability for a given gain requirement.

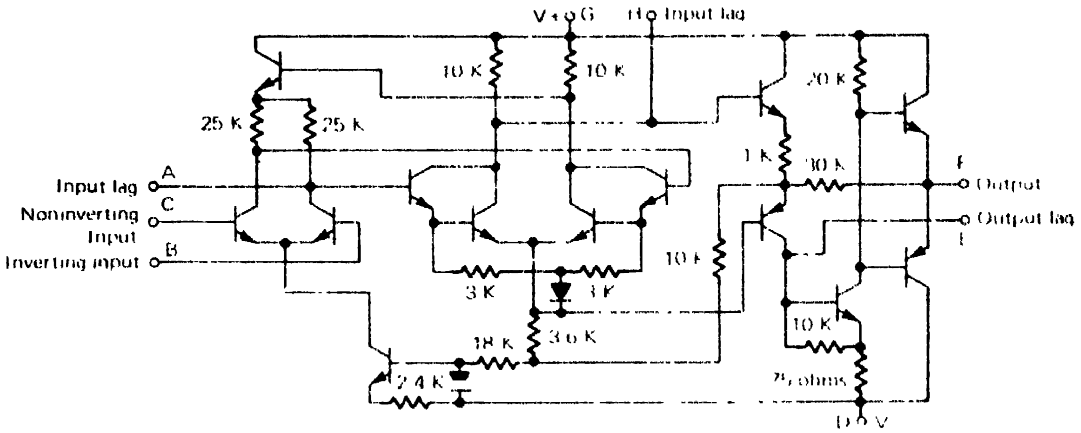


Fig. 23-13. Schematic of a silicon monolithic high-performance operational amplifier. Components shown as diodes are obtained by shorting a transistor's collector to its base and utilizing the base-emitter junction. Type SN52/709. (Courtesy of Texas Instruments Incorporated.)

23-11. INCREASED DIFFERENTIAL INPUT IMPEDANCE

It was mentioned that the differential-signal input impedance in the circuit of Fig. 23-11 is only $2R_1$ and may be too low for many applications. The circuit of Fig. 23-14 has an input impedance limited only by the common-mode input impedance and is normally much higher than $2R_1$.

$$V_o = \frac{R_o + R_1}{R_1}(V_2 - V_1). \qquad (23\text{-}19)$$

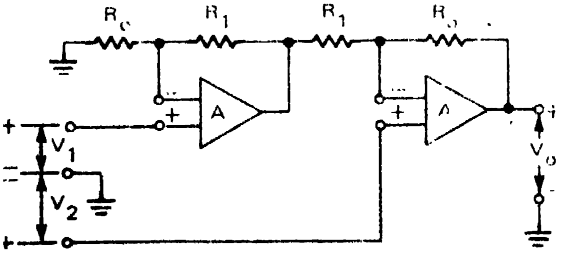


Fig. 23-14. Circuit of a high-differential-input-impedance differential amplifier.

23-12. APPLICATIONS OF DIFFERENTIAL OPERATIONAL AMPLIFIERS

Ground-Noise Rejection

Consider the circuit in Fig. 23-15 to provide a gain of 1000 for the signal source V_1.

The physical distance between source and amplifier means that the noise voltage, V_n, developed in the ground line by ground-loop currents will be significant compared with V_1 and in series with it. Notwithstanding the shielding (grounded at one end only) which reduces magnetic field effects, the output voltage V_o will contain the noise voltage also amplified by 1000. That is,

$$V_o = 1000(V_1 \quad V_n). \qquad (23\text{-}20)$$

The use of a differential amplifier will reduce this error. See Fig. 23-16. Now both inputs to the differential amplifier will have a V_n component that will cancel:

$$\begin{aligned} 1000\Delta V &= 1000[(0 + V_n) - (V_1 + V_n)] \\ &= -1000V_1. \qquad (23\text{-}21) \end{aligned}$$

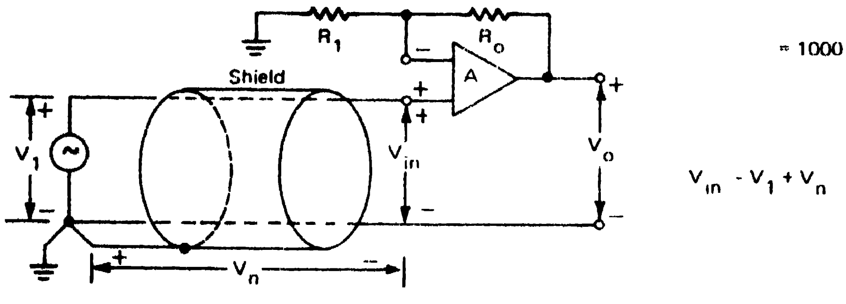


Fig. 23-15. Circuit of a differential operational amplifier which amplifies both the desired signal V_1 and the noise V_n.

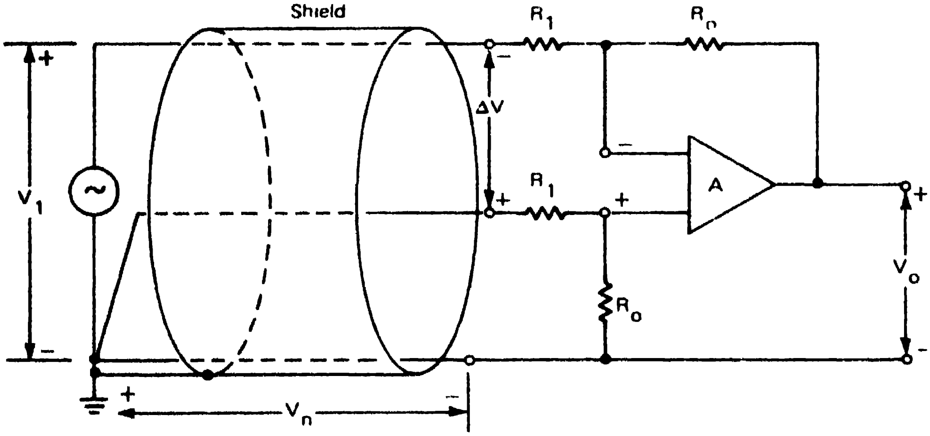


Fig. 23-16. Differential input connection causes V_n to be a common-mode signal that is rejected by the differential amplifier.

However, this requires an amplifier with a very good common-mode rejection; otherwise, the error term introduced due to common-mode gain (see Eq. 23-15) may be comparable to the amplified noise voltage in Fig. 23-15.

Bridge Circuit Measurements

The differential amplifier is a "natural" when it comes to amplifying the output from a Wheatstone bridge circuit, since this is truly a differential signal. Refer to Fig. 23-17.

The Wheatstone bridge may have any one of many transducers in one (or more) of the arms of the circuit. A transducer is a device that converts a physical variable being measured to an electrical signal and is often a resistor that varies with temperature, pressure, etc., such as a thermistor, strain gauge, platinum wire resistance temperature bulb or photocell. In other cases, a voltage may be generated directly.

In the circuit of Fig. 23-17, let T be a thermistor whose resistance is 500 ohms at

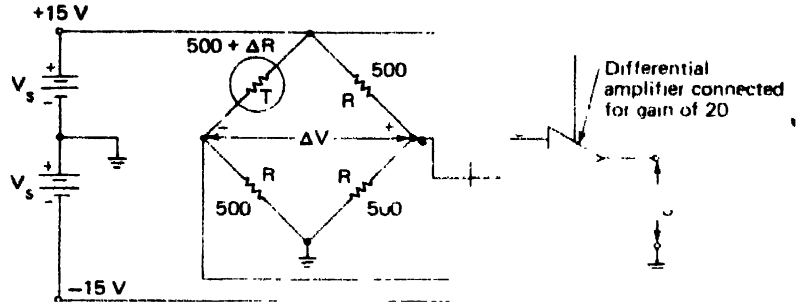


Fig. 23-17. Showing how a differential operational amplifier may be used to amplify the output of a Wheatstone bridge.

25° C and whose change in resistance with temperature is ΔR. With a constant voltage applied to the bridge, the output is

$$\Delta V = \frac{\Delta R}{4R}\,\frac{1}{1+\frac{\Delta R}{2R}} \qquad (23\text{-}22)$$

$$\text{If } \Delta R << 2R, \text{ then } \Delta V \simeq \frac{\Delta R}{4R} V_s. \qquad (23\text{-}23)$$

However, the variation in R, ΔR, is large for most thermistors, as shown in Fig. 23-18(a). Thus, the output voltage is not linear with ΔR or temperature, as shown by the curve in Fig. 23-18(b).

The nonlinearity of ΔV with temperature is due to the nonlinearity of the thermistor and bridge combined.

The purpose of the amplifier is to provide a usable signal that can drive some indicating meter properly calibrated for the units being measured. The amplifier's high input impedance (typically 50 M ohms for an instrumentation amplifier) will not load down the bridge and will not introduce additional nonlinearity, as would a milliammeter connected directly across ΔV. Where necessary, a linearizing circuit may be added to the output of the amplifier to give a straight line relationship between voltage and temperature from -25° C to +50° C, with an accuracy of ±0.2%.

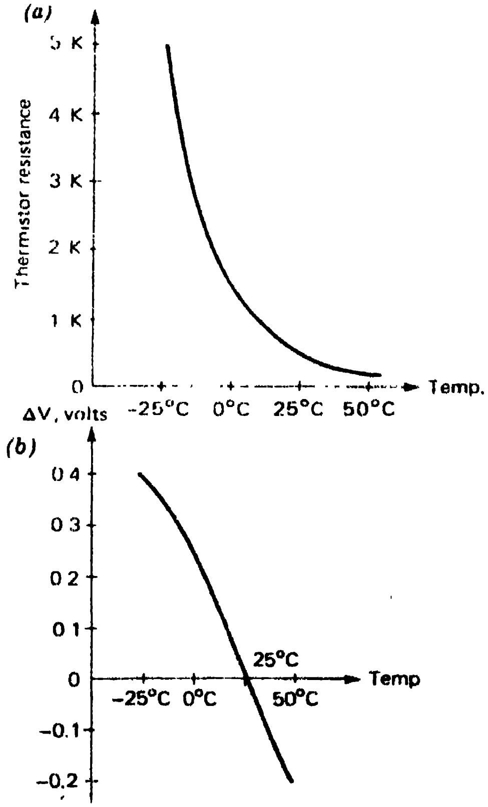


Fig. 23-18. (a) Graph showing the variation of thermistor resistance with temperature. (b) Curve showing the bridge output voltage as a function of temperature. (Courtesy of Burr-Brown Research Corporation.)

Voltage Comparator

It is frequently necessary to provide a signal to indicate when a variable voltage exceeds a certain reference level. For example, it may be required to know when the temperature (in the circuit of Fig. 23-17) exceeds a certain level, so that corrective measures may be taken if necessary. A circuit to do this is shown in Fig. 23-19.

Say that it is required to know when the temperature falls below 0° C. From the preceding curve of ΔV against temperature, $\Delta V = 0.25$ volt @ 0° C, so that V_o in Fig. 23-17 would be 5 volts. In the comparator circuit, $-V_{ref}$ would be -5 volts. As V_{in} (representing the temperature) approaches +5 volts, the voltage at the summing point approaches zero. Say $V_{in} = +4.9$ volts; then the summing point will be negative, and the amplifier with its gain of several thousand will make V_o go positive and tend to saturate at its maximum output (+13 V usually). However, Zener Z1 and diode D1 will conduct, feeding back some voltage to the input to prevent saturation. The output V_o will be limited to $(V_{Z1} + V_{D1})$. (Remember the summing point is a virtual ground.) When $V_{in} = V_{ref}$, the summing point will be zero, and V_o will be zero. (The bridge is balanced and lets the amplifier operate as an open loop amplifier.) If V_{in} exceeds 5 V, the summing point becomes positive, and $V_o = -(V_{Z2} + V_{D2})$. Ideally, the output should "snap" from one state to the other as V_{ref} is reached.

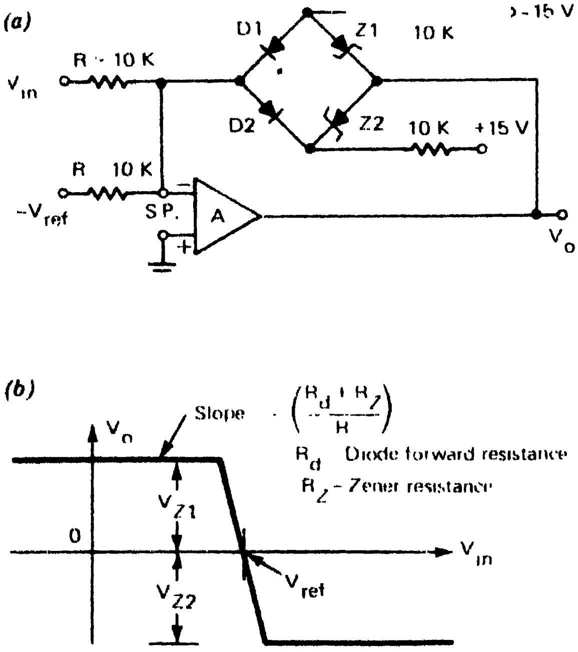


Fig. 23-19. (a) Circuit of a voltage comparator utilizing an operational amplifier. (b) Graph showing the voltage comparator's output as a function of the input. (Courtesy of Burr-Brown Research Corporation.)

In practice, some hysteresis will result, which decreases with increasing gain of the operational amplifier. The point of avoiding saturation of the amplifier is to avoid the excessive "recovery" time required to come out of saturation and change state to the opposite polarity. This time may amount to several seconds in some amplifiers.

Wien Bridge Oscillator

The oscillator shown in Fig. 23-20 was mentioned in Chapter 18 as an example of the use of positive feedback which is at its maximum when

$$f = \frac{1}{2\pi RC} \qquad (23\text{-}24)$$

Then the voltage at A is one-third of V_o and in phase with it. Negative feedback is controlled by R_f and the nonlinear resistance of the lamp which serves to stabilize the amplitude of the oscillations. The output impedance is quite high and should be followed by a buffer amplifier.

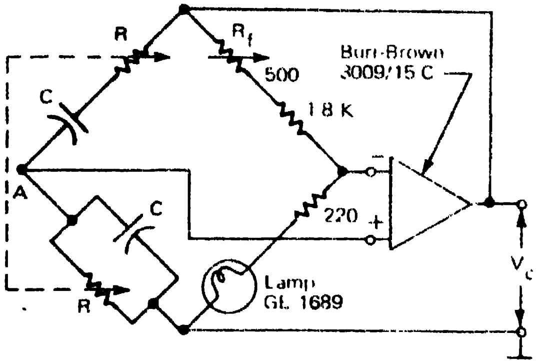


Fig. 23-20. Circuit of a Wien bridge oscillator with a differential amplifier. (Courtesy of Burr-Brown Research Corporation.)

Voltage Regulators

Before examining an operational-amplifier voltage regulator, let us briefly examine the two most common discrete regulators.

Transistorized Shunt Regulator

Referring to Fig. 23-21, the output voltage is determined by the voltage drop across R_s. If I_L should increase due to a load change, V_o will tend to decrease. However, the voltage across R_2 will also decrease, reducing forward bias on the transistor and tending to turn it off. This results in less current flow through the transistor and keeps the voltage drop across R_s relatively unchanged. Thus V_o remains substantially constant. The drawback to this circuit is the large amount of power lost in R_s.

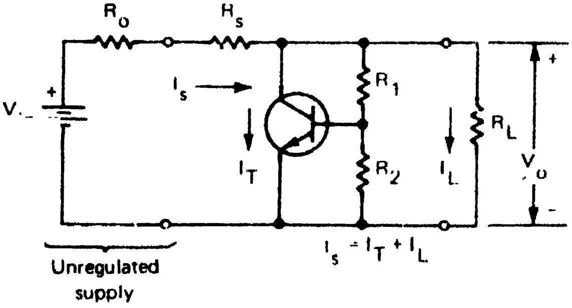


Fig. 23-21. Circuit of a transistorized shunt voltage regulator.

Transistorized Series Regulator

If R_s is replaced by a transistor, a more efficient circuit results that is more sensitive to voltage changes and provides better regulation. See Fig. 23-22.

Q2 actually serves as a differential amplifier where some fraction of the output voltage kV_o is compared with a reference voltage V_R. The difference is amplified by Q2 and appears at the base of Q1; this in turn determines the voltage drop that will occur across Q1. Because of the gain of Q2, it requires only a small change in V_o to have a large effect on Q1. In addition, the output voltage may be varied over a wide range using R_2. The Zener diode and Q2 can be chosen so that the temperature coefficients practically cancel. In fact, an integrated circuit package consisting of a reference diode and NPN transistor on a single chip is available for this purpose.

It should be noted that if R_2 is adjusted for a lower output voltage a greater voltage drop occurs across Q1. Maximum dissipation in Q1 thus takes place at high load currents and low output voltage in variable regulated power supplies employing a series regulator.

Differential-Operational-Amplifier Voltage Regulator

Most operational amplifiers will operate from supplies of between ±12 V to ±18 V with little variation in performance. Thus V_{in} in the diagram of Fig. 23-23 may be anywhere from +24 V to +36 V with the negative supply terminal grounded.

A reference voltage is provided by R_Z and V_Z and is used to drive a constant load of $R_2 + R_3$. C_1 provides filtering of any ripple. The voltage appearing at the noninverting input is then

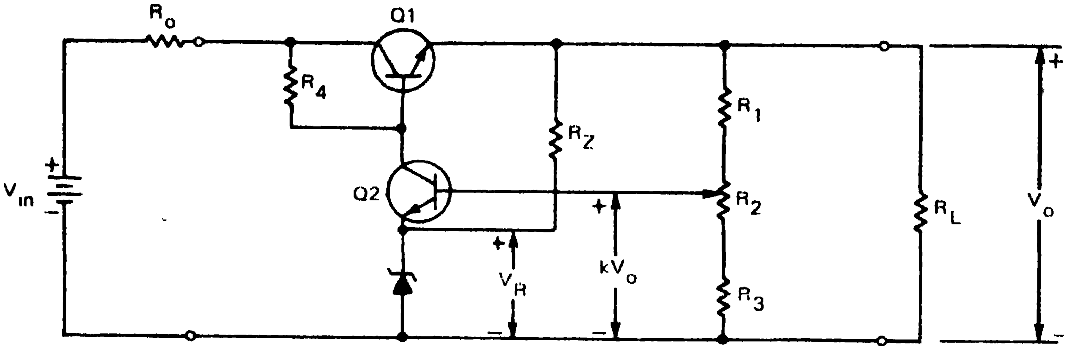

R
o
Q1
R
4
R
Z
R
1
R
2
R
3
R
L
V
in
Q2
V
R
kV
o
V
o
+
−

Fig. 23-22. Circuit of a transistorized series voltage regulator.

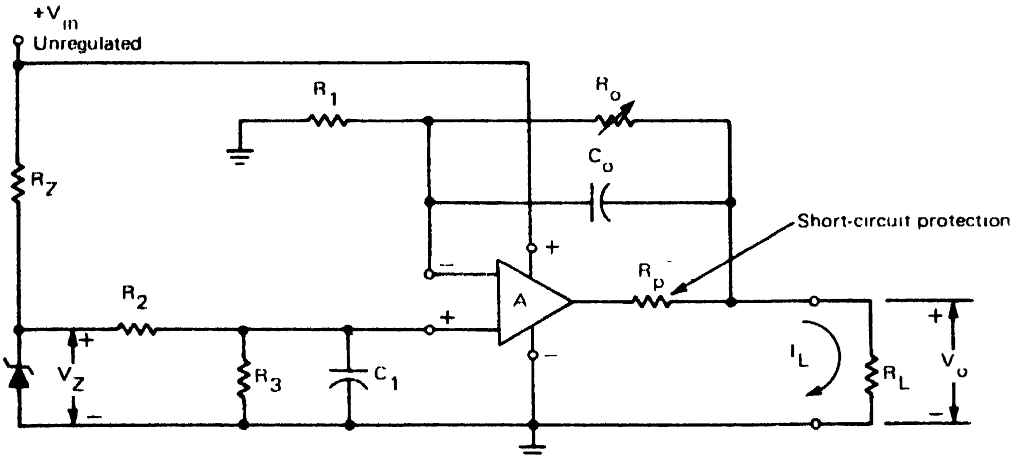

+V
in
Unregulated
R
1
R
o
C
o
R
Z
Short-circuit protection
R
p
R
2
A
V
Z
R
3
C
1
I
L
R
L
V
o
+
−

Fig. 23-23. Voltage regulator using an IC differential operational amplifier. (Courtesy of Burr-Brown Research Corporation.)

$$\frac{R_3}{R_2 + R_3} V_Z. \qquad (23\text{-}25)$$

This voltage is amplified by a factor of

$\dfrac{R_o + R_1}{R_1}$, to yield

$$\frac{R_o + R_1}{R_1} \quad \frac{R_3}{R_2 + R_3} \times V_Z. \qquad (23\text{-}26)$$

Since this equation does not contain V_{in}, theoretically V_o will be independent of V_{in} and I_L within the rated output. In practice load regulation will depend upon the choice of V_Z.

The output voltage is easily varied using R_o, and V_Z must be chosen so that it is in the range of V_o.

It should be mentioned that monolithic voltage regulators are available that can deliver 1 A at 40 V, and with an external power transistor added this can be increased to 10 A of load current.

Other Uses of the Differential Operational Amplifier

At least fifty additional applications of the differential operational amplifier have been documented, so the above selection is very limited. However, it represents a good cross-section of what can be done, which is far more than the name "operational amplifier" once implied. The IC "op-amp" is truly a versatile building block.

23-13. IC AUDIO POWER AMPLIFIER

To show the application of an IC in an audio amplifier and how it is necessary to add many components external to the IC itself, consider the CA 3020 power amplifier in Fig. 23-24, complete with its schematic diagram.

The operation of the CA 3020 can best be seen by considering its functional block diagram, shown in Fig. 23-25.

Circuit Description

Transistor Q1 with the external 5.1 K ohm resistor constitutes an emitter follower to provide an input resistance of 55 K ohms. The output of the buffer amplifier drives a differential amplifier which, with one base grounded (for ac), performs a phase-splitting function to provide two signals 180° out of phase with each other at A and B (see Fig. 23-24(b)). The voltage regulator, with the three diodes constructed from transistor base-emitter junctions and resistor R_{11}, provides the differential amplifier with a collector supply voltage of about 2.1 volts at C and a base bias supply voltage of about 1.4 volts at D.

The driver stage consists of two emitter followers, Q4 and Q5, whose collectors are supplied with the full 9 V. Their outputs drive the power transistors, Q6 and Q7, connected in push-pull to a transformer exterior to the IC. Negative feedback from the outputs of Q4 and Q5 to the bases of transistors Q2 and Q3 through R_5 and R_7, respectively, is provided to help maintain the quiescent or idling currents through the driver and output transistors balanced and constant in the face of any imbalance that may occur in the characteristics of the transistors or their loads. The balance may be enhanced by the inclusion of emitter resistors between terminals 5 and 6 and ground.

When properly heat-sinked, the above audio amplifier will deliver one-half watt into a 3.2-ohm speaker with 35 mV applied. Since the CA 3020 has an 8 MHz bandwidth, the possibility of oscillation can be minimized by restricting the overall bandwidth by adding bypass capacitor C_4 and following good high-frequency wiring practices. The low frequency response may be increased by increasing the values of C_1, C_2, and C_3, while C_4 is primarily responsible for establishing the upper frequency limit.

23-14. SUMMARY

(1) A differential operational amplifier has inverting and noninverting inputs with high input impedance and differential or open-loop gains between 1,000 and 10 million.

(2) When the inverting input is used with negative feedback due to R_o, the closed loop gain is given by $-R_o/R_1$, and the input impedance is R_1; the output impedance is the open-loop output impedance divided by the loop gain.

(3)

$$\text{Loop Gain} = \frac{\text{Open Loop (or Differential) Gain}}{\text{Closed Loop Gain}}$$

or, Loop Gain (dB) = Open Loop Gain (dB) - Closed Loop Gain (dB).

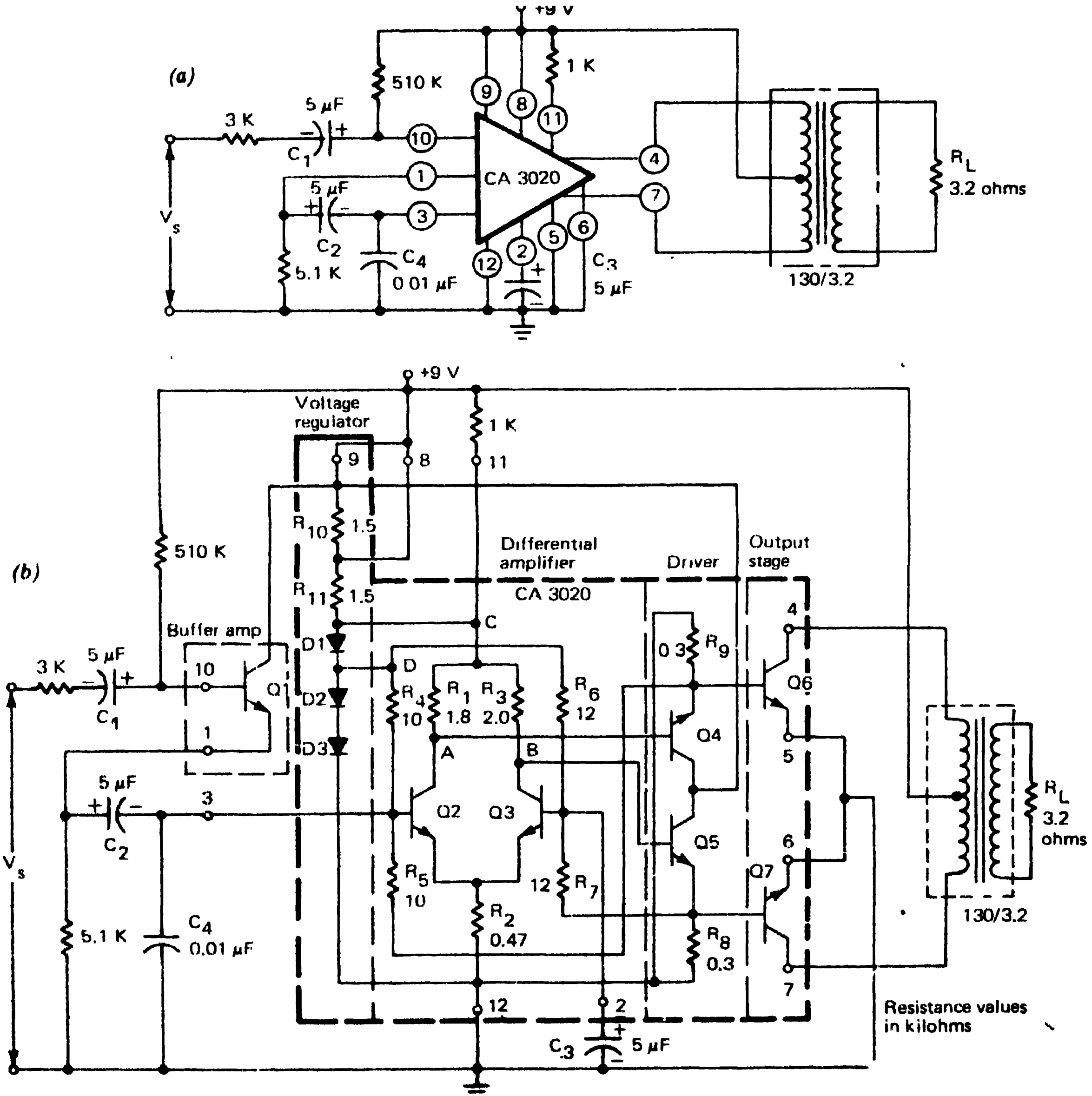

(a)
+9 V
510 K
1 K
3 K
5 μF
C_1
CA 3020
5 μF
C_2
5.1 K
C_4
0.01 μF
C_3
5 μF
V_s
R_L
3.2 ohms
130/3.2
(b)
+9 V
Voltage regulator
1 K
510 K
R_{10} 1.5
R_{11} 1.5
D1
D2
D3
Buffer amp
Q1
Differential amplifier
CA 3020
Driver
Output stage
C
D
R_4 10
R_1 1.8
R_3 2.0
R_6 12
A
B
Q2
Q3
R_5 10
R_2 0.47
12 R_7
0.3 R_9
Q4
Q5
Q6
Q7
R_8 0.3
3 K
5 μF
C_1
5 μF
C_2
5.1 K
C_4
0.01 μF
C_3
5 μF
V_s
R_L
3.2 ohms
130/3.2
Resistance values in kilohms

Fig. 23-24. (a) Circuit of an audio amplifier utilizing an IC that can deliver 1/2 watt to R_L. (b) Schematic diagram of the CA 3020 IC audio amplifier, showing the differential amplifier with the base of Q3 grounded for ac signals through C_3. (Courtesy of RCA Corp.)

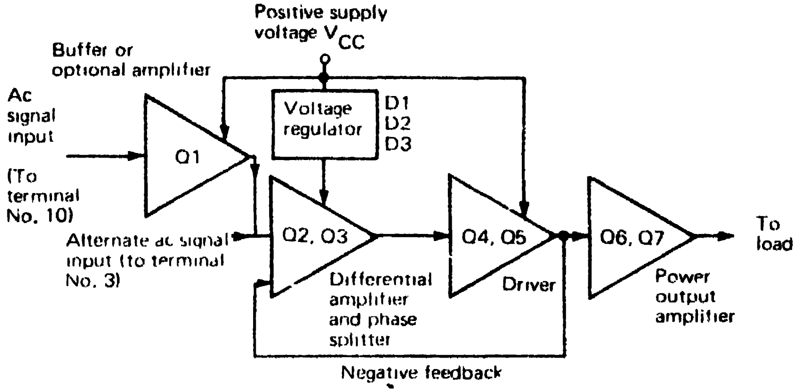

Positive supply voltage V_{CC}
Buffer or optional amplifier
Ac signal input
(To terminal No. 10)
Q1
Voltage regulator
D1
D2
D3
Alternate ac signal input (to terminal No. 3)
Q2, Q3
Differential amplifier and phase splitter
Q4, Q5
Driver
Q6, Q7
Power output amplifier
To load
Negative feedback

Fig. 23-25. Block diagram of the CA 3020 IC audio power amplifier. (Courtesy of RCA Corp.)

(4) If a number of signals are connected to the inverting input, the output is proportional to the sum of the input signals, making possible the solution of linear algebraic and differential equations on an analog computer.

(5) One way to increase input impedance to an amplifier utilizing the inverting input is to reduce the feedback voltage by connecting a voltage divider at the output.

(6) If the noninverting input to a differential operational amplifier is used, the input impedance is increased to a value given by

$$R_{cm} \parallel (R_i \times \text{Loop Gain}),$$

where R_{cm} is the common mode impedance and R_i is the differential impedance.

(7) The voltage gain for a noninverting input is given by

$$\frac{R_1 + R_o}{R_1}$$

(8) A true ideal differential amplifier amplifies the difference between two input voltages providing an output given by

$$V_o = \frac{R_o}{R_1} \Delta V, \text{ where } \Delta V = V_2 - V_1.$$

(9) In practice, the output also consists of an error term that is due to the common mode input voltage,

$$= \frac{V_1 + V_2}{2},$$ such that the total output of a differential amplifier is given by

$$V_o = \frac{R_o}{R_1} \Delta V + \frac{R_o}{R_1} V_{cm}' \frac{1}{\text{CMRR}},$$

where $V_{cm}' = V_2 \dfrac{R_o}{R_o + R_1}$, and CMRR = the common mode rejection ratio.

(10) The common mode rejection ratio is a figure of merit of a differential amplifier, since it is the ratio of differential gain, A_d (the desired gain), to the common-mode gain, A_c.

$$\text{CMRR} = \frac{A_d}{A_c} \text{ or } \text{CMR}_{dB} = 20 \log A_d - 20 \log A_c.$$

(11) Emitter-coupled differential amplifiers are the type of circuit used predominantly in ICs, because of the manufacturing ability to closely match components, and since the devices are so closely spaced their variations due to temperature tend to cancel, providing excellent dc coupling stability.

(12) High values of CMRR are provided by large effective values of emitter resistance, R_e, in the emitter-coupled amplifier. This may be provided by a transistor or CRD to provide essentially a constant-current source in the emitter. Thus

$$\text{CMRR} \approx \frac{h_{fe}R_e}{h_{ie}}$$

(13) An excellent application of a differential operational amplifier is the rejection of ground noise picked up by the cable connecting signal to amplifier.

(14) An IC audio power amplifier also utilizes a differential amplifier, where, with one base grounded for ac, a phase-splitting function is provided to drive push-pull outputs and supply power to an externally connected transformer.

QUESTIONS

Q23-1. (a) What was the operational amplifier originally used for?

(b) How does it differ from present differential operational amplifiers?

(c) What is the range of open-loop dc gain found in operational amplifiers?

(d) How does the closed-loop voltage gain differ from the open-loop gain?

Q23-2. (a) Referring to Fig. 23-2, which point may be thought of as a virtual ground? Why?

(b) What does this imply about the input resistance seen by V_1?

(c) What is meant by the loop gain of the circuit?

(d) How does this affect the closed loop output impedance?

Q23-3. (a) Which is the summing point in Fig. 23-3?

(b) What does this mean?

(c) If V_1, V_2, and V_3 were all negative would the output be positive or negative?

Q23-4. If an operational amplifier has a maximum output swing of ± 10 V, what must we be sure to avoid in using this type of amplifier in an analog computer to solve linear equations?

Q23-5. (a) What difficulties arise when the input impedance to an inverting amplifier is increased by increasing the size of feedback resistors?

(b) What is one way of obtaining an increase in input impedance without sacrificing gain?

Q23-6. (a) How does the closed-loop voltage gain in a noninverting amplifier compare with that in an inverting amplifier for the same values of resistors?

(b) How does the input impedance compare? What is the expression for this impedance?

Q23-7. (a) What is the purpose of a differential amplifier?

(b) What is meant by the common-mode input voltage?

(c) What is the difference between the common-mode gain and the differential gain? What would be *ideal* values for each?

Q23-8. (a) How is the ability of an amplifier to provide gain for the differential signal but reject the common-mode signal indicated?

(b) What is this called when expressed in decibels?

(c) What is the expression that gives the output voltage from a true differential amplifier?

(d) How may the error term be minimized *before* ordering the amplifier and also *after* the amplifier has been chosen?

Q23-9. (a) A successful emitter-coupled differential amplifier operates with essentially a constant current source in its emitter lead: true or false?

(b) What three ways may be used to provide the operation in (a)?

(c) Which is the most successful and why?

(d) What important parameter benefits from the use of as high a value of effective emitter resistance as possible?

Q23-10. (a) What type of connection is used in Fig. 23-15, inverting or noninverting?

(b) What does the input to the noninverting terminal consist of?

(c) What type of connection is used in Fig. 23-16?

(d) What does the input voltage to the R_1 resistors consist of?

(e) What is the common-mode voltage?

(f) Under what conditions will the circuit of Fig. 23-15 be successful in rejecting ground noise?

Q23-11. (a) What voltage is applied to the bridge circuit in Fig. 23-17?

(b) If the thermistor's resistance is 500 ohms, what are ΔV and V_o?

(c) If an increase in temperature occurs, does the current through the thermistor increase or decrease?

(d) Is ΔV positive? Is V_o positive?

Q23-12. Assume the input to the voltage comparator in Fig. 23-19(a) is connected to the output of the amplifier in Fig. 23-17. It is required to know when the temperature exceeds 50°C.

(a) When the temperature is 50°C, what is the output, V_o', in Fig. 23-17?

(b) What must be the value of $-V_{ref}$ in Fig. 23-19(a)?

(c) What will be the output of the voltage comparator when the temperature is 25°C?

(d) Describe what takes place as the temperature approaches 50°C and just exceeds it.

(e) Draw a graph of the comparator's output voltage for the conditions in (d)

Q23-13. (a) Explain in your own words the voltage-regulating action of the shunt and series regulators in Figs. 23-21 and -22.

(b) What is the main disadvantage of the shunt regulator?

(c) When does maximum dissipation occur in the transistor in the shunt regulator, at maximum or minimum load?

(d) Why is the series regulator more sensitive to a voltage change?

(e) When does maximum dissipation occur in the series power transistor, Q1?

Q23-14. Which equation suggests that the voltage regulation of the circuit in Fig. 23-23 should be very good? Explain.

Q23-15. (a) Explain how the differential amplifier in the CA 3020 of Fig. 23-24(b) functions as a phase splitter.

(b) If the input signal to the amplifier were connected to the alternate input at terminal 3, what would be the major difference as far as the input signal is concerned?

(c) Why do you think it is necessary to connect so many components exterior to the IC itself?

PROBLEMS

P23-1. An operational amplifier with an open loop gain of 100,000 is used in the inverting mode as in Fig. 23-2. R_o = 1 M ohm and R_1 = 100 K ohms. If the amplifier's open loop output impedance is 10 K ohms, determine:

(a) The voltage gain, V_o/V_1.

(b) The input impedance seen by V_1.

(c) The loop gain as a ratio and in decibels.
(d) The overall output impedance.

P23-2. Repeat P23-1 using R_1 = 10 K ohms.

P23-3. The summing amplifier in Fig. 23-3 has R_o = 10 K ohms, R_1 = 10 K ohms, R_2 = 2.2 K ohms and R_3 = 3.3 K ohms. If V_1 = +6 V, V_2 = -3 V and V_3 = -0.75 V what is V_o?

P23-4. (a) Solve Eq. 23-1 for V_o/V_1 and compare with Eq. 23-6.

(b) A' is the effective open-loop gain taking into account the loading due to $R_o \| R_1$. When will it be substantially the same as A?

P23-5. Draw an analog computer circuit to solve the following equations:

$$X - 4Y = 6,$$
$$3X + 2Y = 5$$

Indicate possible values of fixed standard resistors assuming ±5% tolerances are available. Indicate where voltmeters should be connected to read the solutions and show what they are.

P23-6. If in Fig. 23-4 the 2 K ohm resistor were replaced by a 3.3 K ohm resistor and the 3 K ohm resistor by a 2.2 K ohm resistor, what two equations would be wired for solution?

P23-7. Calculate the gain and input impedance of the amplifier in Fig. 23-6 with R_o = 10 M ohms, R_1 = 100 K ohms, and R_i equal to:

(a) 0.1 M ohm,
(b) 1 M ohm; A = 5,000.

P23-8. In Chapter 17 we showed that

$$A_{vf} = \frac{A_v}{1 - kA_v}.$$

Use the equation, $A_{vf} = R_o/R_1$, to determine what the feedback factor, k, is in a circuit such as that in Fig. 23-6. Hence show that the circuit of Fig. 23-7 will give a gain of 1000 if R_1 = 100 K ohms and R_o = 1 M ohm.

P23-9. An economy IC amplifier is used in the noninverting mode as in Fig. 23-8, with R_o = 10 M ohms and R_1 = 1 M ohm. If A = 20,000, R_{cm} = 200 M ohms, R_i = 300 K ohms, and $R_{o,OL}$ = 5 K ohms, calculate:

(a) Voltage gain.
(b) Input impedance.
(c) Output impedance.
(d) The new voltage gain, and input and output impedances if R_1 is changed to 100 K ohms.

P23-10. Repeat P23-9 if an IC amplifier with FET inputs is used with A = 400,000, $R_i = R_{cm} = 10^{12}$ ohms, and $R_{o,OL}$ = 1 K ohm.

P23-11. An IC operational amplifier with a differential open-loop voltage gain of 20,000 has a common-mode rejection of 80 dB.

(a) What is its CMRR?
(b) What is its common-mode gain?
(c) What is the output from the amplifier if, in Fig. 23-9, the input at B is 0.18 mV and at A is 0.16 mV?
(d) Repeat (c) for 2.00018 V and 2.00016 V, respectively.

P23-12. An FET-input IC operational amplifier has a typical open-loop differential gain of 100,000 and a common-mode gain of 25.

(a) What is its CMRR?
(b) What is its CMR in decibels?
(c) What is the output from the amplifier if in Fig. 23-9, the input at B is 0.18 mV and at A is 0.16 mV?
(d) Repeat (c) for 2.00018 V and 2.00016 V, respectively. (The amplifier saturates at ±10 V.)

P23-13. Consider the differential amplifier in Fig. 23-11 with $R_o = 100$ K ohms and $R_1 = 10$ K ohms. If the CMRR = 100, determine the output with:

(a) V_2 2 V and $V_1 = 1$ V.
(b) $V_2 = 1$ V and $V_1 = 2$ V.
(c) Repeat (a) and (b) if the CMRR is increased to 1000.
(d) What is the differential input impedance?

P23-14. The CMRR of the amplifier in Fig. 23-10 is being determined experimentally. With the same voltage of 1 V applied to both A and B, the output is 2.5 V. When B is grounded and 1 mV is applied to A, the output is 3 V.

(a) What is the CMRR?
(b) If this amplifier is used in the circuit of Fig. 23-11 with $R_o = 100$ K ohms and $R_1 = 1$ K ohm, calculate the output voltage when:
    (i) $V_2 = 9.00$ V and $V_1 = 9.05$ V.
    (ii) $V_2 = 1.00$ V and $V_1 = 1.05$ V.

P23-15. Refer to Fig. 23-12. Let $R_{L1} = R_{L2} = 10$ K ohms in all three circuits, $h_{ie} = 1$ K ohm, $h_{fe} = 100$. For each circuit calculate:

(i) A_c (ii) A_d (iii) CMRR, given,
(a) $R_e = 2.2$ K ohms in Fig. 23-12(a).
(b) Effective value of $R_e = 47$ K ohms in Fig. 23-12(b).
(c) Effective value of $R_e = 1$ M ohm in Fig. 23-12(c).

P23-16. (a) In Fig. 23-15, V_1 is a 0.5 V p-p signal at 1 kHz; $R_o = 100$ K ohms, $R_1 = 10$ K ohms. If V_n is 0.1 V p-p at 60 Hz, what does the output consist of?

(b) If the circuit is modified to that in Fig. 23-16 (same value of R_o and R_1), what will be the output under ideal conditions (infinite CMRR)?
(c) Repeat (b) taking into account the error term due to a finite CMRR of (i) 10, (ii) 1000.
(d) Repeat (c) with the inputs to the inverting and noninverting terminals interchanged (V_1 to plus and V_n to minus).

24

AN INTRODUCTION TO THE FABRICATION OF INTEGRATED CIRCUITS

24-0. INTRODUCTION

Integrated circuits form a branch of the general field of microelectronics which originated in the late 1950s. At that time the military was interested in reducing the size of its electronic equipment to approximately one tenth of its initial volume. However what started as a drive for miniaturization was soon tempered by goals of reliability and cost, until today the prime advantage of microelectronics (integrated circuits) over discrete components is the increased reliability. In a survey conducted in late 1969 among control engineers on the reasons for using integrated circuits for switching applications, increased reliability was cited as the most important reason, followed by increased system complexity, more economical approach, increased speed requirements and finally, smaller physical size.

24-1. EVOLUTION OF INTEGRATED CIRCUITS

Integrated circuits evolved from the miniaturization program. The first effort was to miniaturize discrete components. This was followed by developing printed circuit boards (single- and double-sided) to eliminate bulky wiring and tie points; discrete components were then placed on the board. Conserving space by packaging innovations led to the use of modules in which discrete components were placed like "cordwood" between planes of printed circuit boards. This was still a 3D (three dimensional) package, and a method to provide 2D or "flat" circuitry was sought. A partial solution was found for passive components, wherein resistors and capacitors were made by depositing thin films of metal on a substrate and adding discrete active devices afterward. Interconnections were then a problem in complex circuits, which was partially overcome by "burrowing" into the substrate and making connections by a tunnel under other parts and connections. This substrate was a passive substance such as pyrex. However, with the advent of making many transistors on a single substrate by the diffusion process, it was a short step for interconnections to be made between many different components (passive and active) on the same active substrate, resulting in monolithic integrated circuits as we know them today.

Interconnections of various combinations of thin films, discrete components, and monolithic circuits led to the name of hybrid circuits, which use a mixture of the different processes.

The terminology used in the above discussion is illustrated in the "microelectronics tree" shown in Fig. 24-1.

Here is shown a block called "functional devices," which includes devices such as the piezoelectric quartz crystal, often found in oscillators. It has the property of directly converting a mechanical force to an electrical output, and vice versa. It can be used as an electrical filter to pass only a certain band of frequencies, or it can be used to delay an electrical signal, as in a delay line. In either case, it takes the place of a large number of RLC discrete components and is clearly a contribution to microelectronics.

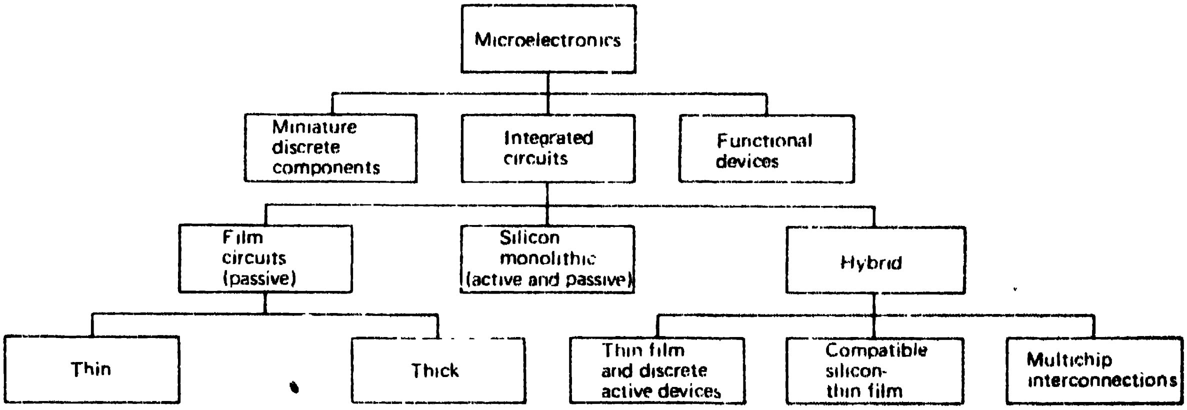

Microelectronics
Miniature discrete components
Integrated circuits
Functional devices
Film circuits (passive)
Silicon monolithic (active and passive)
Hybrid
Thin
Thick
Thin film and discrete active devices
Compatible silicon-thin film
Multichip interconnections

Fig. 24-1. Microelectronics "tree" showing what constitutes the field of microelectronics.

24-2. THE MONOLITHIC INTEGRATED CIRCUIT

Let us consider briefly the steps in fabricating the simple circuit shown in Fig. 24-2(a) on a single chip of silicon.

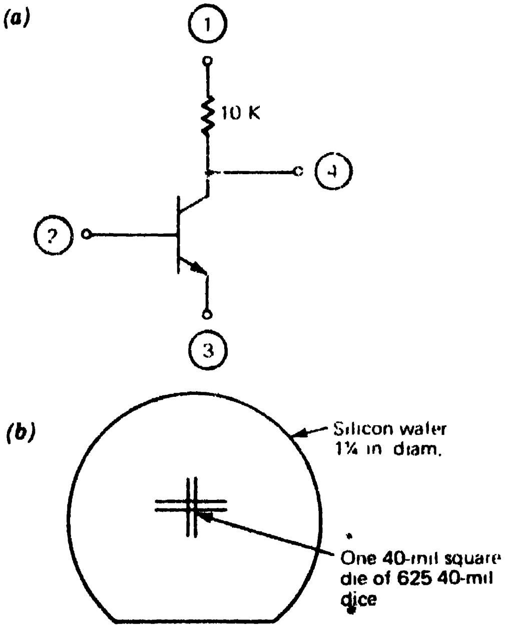

(a)
1
10 K
4
2
3
(b)
Silicon wafer 1¼ in diam.
One 40-mil square die of 625 40-mil dice

Fig. 24-2. (a) Circuit to be constructed on a silicon chip. (b) Silicon wafer that typically contains 625 chips.

We will look at just one die (or chip) on the whole wafer, but it must be emphasized that the steps that follow are simultaneously producing the same circuit on all of the 625 chips in the wafer, which may then be cut up and the resulting dice mounted in separate packages. See Fig. 24-2(b).

Wafer Preparation

A single crystal of extremely pure P-type silicon (typically 6 inches long by $1\frac{1}{4}$ inches in diameter) is sliced into wafers 10 to 15 mils thick. These wafers, after being lapped and polished to a mirror finish, are approximately 5 mils thick. The silicon crystal has a flat ground on one side which is used to index (position) the wafer in succeeding operations. The lightly doped P-type wafer provides the base or substrate on which the transistor and other elements are built.

Epitaxial Growth

An N-type layer of silicon, typically 1 mil thick, is now grown on the P-type substrate by placing the wafer in a furnace at 1200°C and introducing a gas with a phosphorus content (donor material). This provides the collector material for the NPN transistors. A cross-sectional view of *part* of the wafer is shown in Fig. 24-3.

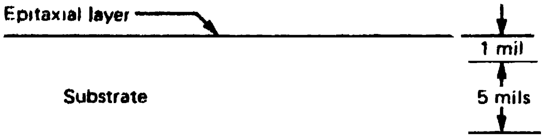
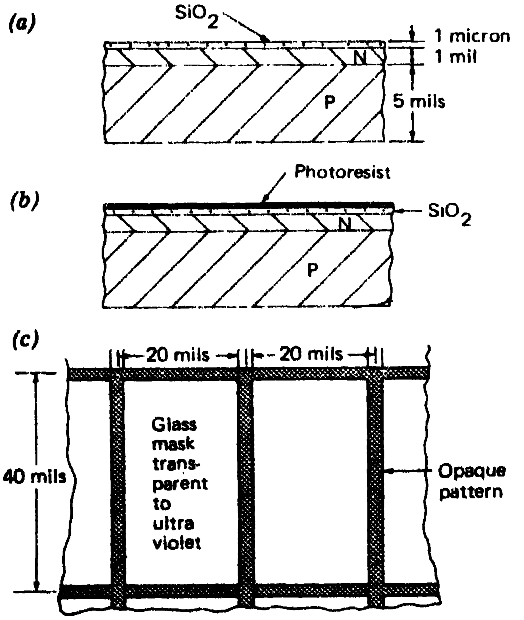

Epitaxial layer
1 mil
Substrate
5 mils

Fig. 24-3. Cross-section of silicon wafer showing the N-type epitaxial layer.

Diffused Isolation

The N-type epitaxial layer is now isolated into islands, so that on each island a transistor or some other component may be fabricated. This is done by diffusing P-type material through the epitaxial layer to the substrate by a series of steps that is common in all diffusion processes.

(a) A thin layer of silicon dioxide (approximately 1 micron thick) is formed on the N-type layer by oxidizing the wafer in dry oxygen. See Fig. 24-4(a).


(a)
SiO2
1 micron
1 mil
N
P
5 mils
(b)
Photoresist
SiO2
N
P
(c)
20 mils
20 mils
40 mils
Glass mask transparent to ultra violet
Opaque pattern

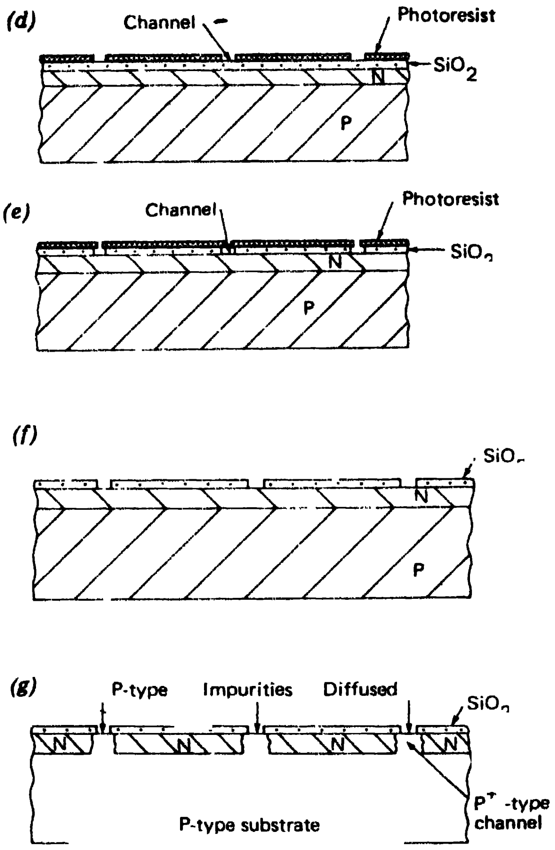

(d)
Channel
Photoresist
SiO2
N
P
(e)
Channel
Photoresist
SiO2
N
P
(f)
SiO2
N
P
(g)
P-type
Impurities
Diffused
SiO2
N
N
N
N
P+ -type channel
P-type substrate

Fig. 24-4. Sequence of steps used in the diffused isolation process (not drawn to scale).
(a) Silicon dioxide layer due to oxidation of the epitaxial layer.
(b) Coating of photoresist that will harden when exposed to ultraviolet.
(c) Portion of a mask that selects the areas to be etched.
(d) Showing the wafer where the photoresist was removed by washing away the unexposed areas selected by the masking process.
(e) Showing how the silicon dioxide may be etched away through the photoresist channels.
(f) Showing the channels after removal of the photoresist.
(g) Showing the effect of diffusing P-type impurities through the channels to provide isolated N-type islands.

(b) A thin uniform coating of a chemical called photoresist is now deposited on the SiO_2 layer. This is a liquid that will harden when exposed to ultraviolet light. See Fig. 24-4(b).

(c) A mask (which is a glass plate with a pattern drawn on it) is now placed over the photoresist surface. The purpose of the mask is to select the areas through which ultraviolet light may be directed to cause a hardening of the photoresist. Conversely, where the dark pattern of the mask comes in contact with the photoresist, no hardening will take place and the photoresist may be washed away by a developer that further hardens the areas that were exposed to the ultraviolet light. A portion of the mask used in our example may appear as shown in Fig. 24-4(c).

The masks are drawn to scale some 200 to 1000 times actual size and then reduced photographically. Many different masks with various patterns are required to develop an integrated circuit. This portion of IC fabrication is referred to as photolithography, and the size of the IC is limited by the photographic resolution possible.

(d) After the wafer and mask have been exposed to ultraviolet light and the unexposed areas washed away, the wafer will have the appearance shown in Fig. 24-4(d).

(e) The SiO_2 coating may then be removed by an etching process from the areas not protected by the photoresist, as indicated in Fig. 24-4(e).

(f) The photoresist is now removed completely by scrubbing with heated solvents. It was used solely for providing selective etching of the SiO_2. The result is shown in Fig. 24-4(f).

We are now ready for the diffusion process, but it should be noted that the above steps using the particular mask needed for the pattern required are repeated *prior to any new diffusion.*

(g) The silicon wafers are now placed in a "boat" and passed through a furnace containing a gaseous boron (acceptor material) atmosphere. The P impurities will diffuse into the silicon (like water being absorbed by a sponge), turning the N-type material into a P type channel down to a depth extending to the P-type substrate. This will result in islands of N-type material resting on the P-type substrate under the SiO_2 layer, as illustrated in Fig. 24-4(g).*

It should now be evident that the purpose of the SiO_2 layer was to select where the P-type impurities should diffuse, maintaining and protecting the N-type islands. Actually the concentration of the P-type channels is somewhat higher than that of the P-type substrate and is often referred to as P^+.

Base Diffusion

The P-type base of the transistor now needs to be diffused into the collector. At the same time, diffusion of the resistor takes place in the adjoining island. First, a complete layer of SiO_2 is provided, after which the photoresist and masking process is repeated, as in steps (a), (b), (c), and (d) above. A portion of the mask used in this case may have the form shown in Fig. 24-5(a).

And a cross-sectional view through X-X of the resulting wafer, after exposure to ultraviolet and subsequent washing, is shown in Fig. 24-5(b).

Following the SiO_2 etching as in step (e) above, and the photoresist removal as in step (f), the result will be exposed N-type areas or "windows," as in Fig. 24-5(c).

---

*Not shown in the above sequence is a method of reducing the resistance of the N-type collector structure. This is achieved by some manufacturers providing N^+-type pockets, or "buried layers," immediately under the N-type collector in the P-type substrate.

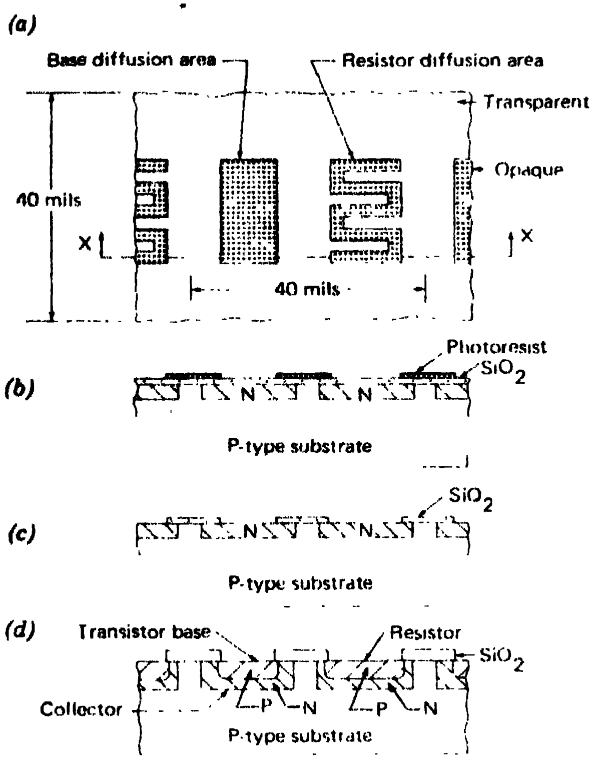

(a)
Base diffusion area
Resistor diffusion area
Transparent
Opaque
40 mils
X
X
40 mils
(b)
Photoresist
SiO2
N
N
P-type substrate
(c)
SiO2
N
N
P-type substrate
(d)
Transistor base
Resistor
SiO2
Collector
P
N
P
N
P-type substrate

Fig. 24-5. Sequence of steps used in the base diffusion process (not to scale).
(a) Portion of the mask to provide the base and resistor diffusion areas.
(b) Cross-sectional view of the resulting wafer through X X, after exposure to ultraviolet and washing.
(c) Showing the N-type windows following etching and photoresist removal.
(d) Showing the diffused P-type transistor base and resistor.

A P-type diffusion in a phosphorus atmosphere will form a base of the required thickness (determined by the length of time diffusion is allowed to take place) and a resistor of the same thickness. See Fig. 24-5(d).

It is to be noted that, while the base is a rectangular diffused area, the resistor will occupy a zigzag diffused area to make up the necessary resistance.

Emitter Diffusion

Following the same procedure as outlined above, the N-type emitter is now diffused into the base following the photoresist and masking process. The result is illustrated in Fig. 24-6. (No diffusion takes place in the adjoining island, since the resistor is complete.)

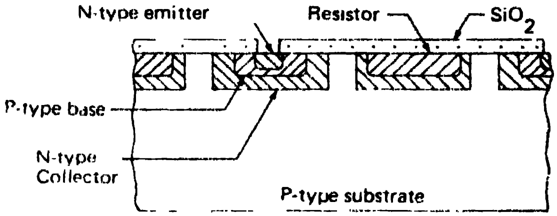

N-type emitter
Resistor
SiO2
P-type base
N-type Collector
P-type substrate

Fig. 24-6. Showing the completed transistor after emitter diffusion.

Pre-ohmic Etch

It then remains to make provision for metal ohmic (nonrectifying) contacts with the diffused areas. This is done once again by the SiO_2 layer, photoresist, and masking process to provide access to the layers, as depicted in Fig. 24-7.

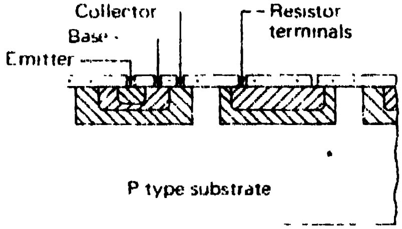

Collector
Base
Emitter
Resistor terminals
P type substrate

Fig. 24-7. Pre-ohmic etch provides access to make connections to the components.

It can be seen that double openings could be made to the collector and base, and this is sometimes done where high current capability (and reduced voltage drop) is required.

Metallization

Ohmic contacts and interconnections are now made, using aluminum for good adhesion to the silicon and silicon dioxide. Bonding pads are provided around the circumference of the die (within the 40 mil square area) for later connection of wires. This metallization is carried out by evaporating aluminum over the entire surface (approximately 1 micron thick) and then selectively etching away the aluminum to leave the desired interconnections and bonding pads. See Fig. 24-8.

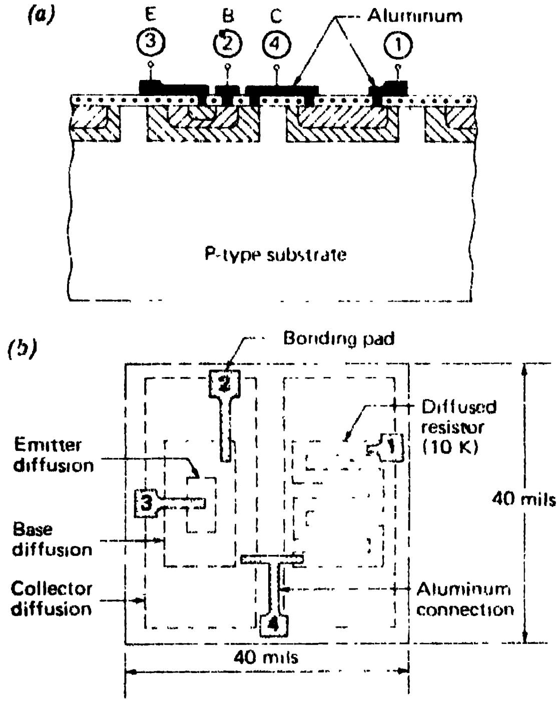


Fig. 24-8. (a) Cross-section through a portion of the wafer showing the metal connections.
(b) Plan view of the chip showing the metallization connections and bonding pads.

Circuit Probing

Each integrated circuit on the wafer is now checked electrically for proper performance characteristics by placing probes on the bonding pads and performing some test. Faulty circuits are marked and discarded after dicing.

Dicing

Dicing is the process of separating the wafer into individual chips or dice by scribing the wafer and breaking.

Mounting and Packaging

The individual die is very small and brittle and must be cemented or soldered to a gold-plated header through which leads are already connected. Many packages exist to hermetically seal the circuit, such as the round TO-5 can, which can accommodate up to 14 pins, or the TO-91 flat package, or the dual in-line-packages. As the complexity of ICs increases, packaging and lead connections become a real problem. Some of the common integrated-circuit packages available are shown in Fig. 24-9.

Wire Bonding

Connections between the IC die and the package leads is done by a thermal compression method using either aluminum or gold wire (under a microscope).

A view of "our" integrated circuit mounted in a TO-5 can without the cap is shown in Fig. 24-10.

It can be seen that the size of the die is quite small compared with the header. Dice may vary from 0.020 inch square up to 0.2 inch square, depending upon circuit complexity. Large scale integration (LSI) involves larger sizes and will be referred to again in Section 24-6.

Encapsulation

A cap is now placed over the circuit, and sealing takes place in an inert atmosphere.

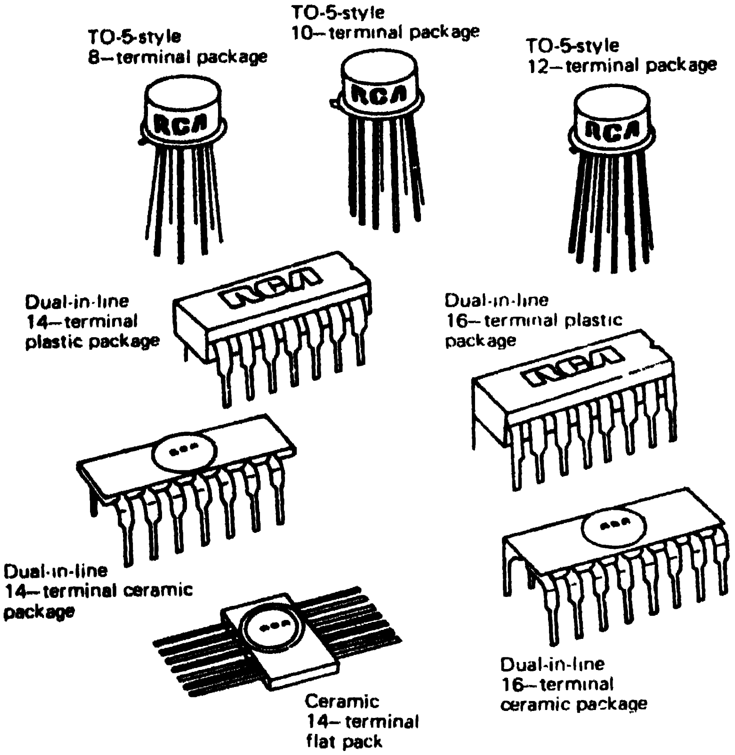

TO-5-style
8—terminal package
TO-5-style
10—terminal package
TO-5-style
12—terminal package
RCA
Dual-in-line
14—terminal
plastic package
Dual-in-line
16—terminal plastic
package
Dual-in-line
14—terminal ceramic
package
Ceramic
14—terminal
flat pack
Dual-in-line
16—terminal
ceramic package

Fig. 24-9. Showing some of the integrated-circuit packages available. (Courtesy of the RCA Corp.)

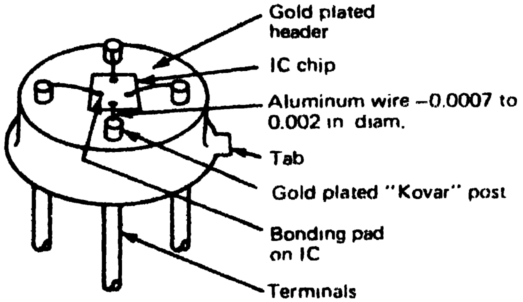

Gold plated
header
IC chip
Aluminum wire –0.0007 to
0.002 in diam.
Tab
Gold plated "Kovar" post
Bonding pad
on IC
Terminals

Fig. 24-10. Showing the mounted IC chip before encapsulation.

Final Testing of the circuit is now carried out.

It takes typically about 100 hours and 35 major steps to process a silicon wafer, although of course as many as 200 wafers may go through a diffusion process at one time.

It is interesting to note that, to obtain a total yield of 30% from the completed wafer (188 satisfactory circuits from a possible 625), each step requires an average accuracy of 97%.

24-3. INTEGRATED CIRCUIT COMPONENTS

Let us summarize the different methods of producing active and passive components and some of their major characteristics.

Resistors

Resistors are obtained either by utilizing the resistivity of diffused areas (usually P-type as in the above example) or films of metal laid on a substrate. In either case, a term called "sheet resistance" is useful in describing the values obtainable. Consider the resistance of the two squares of equal thickness shown in Fig. 24-11.

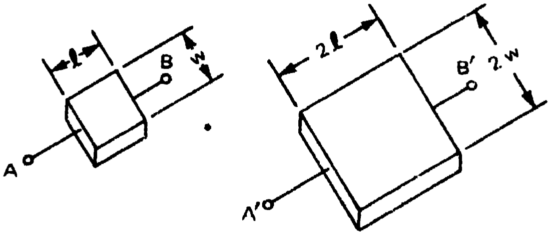


Fig. 24-11. Two squares of equal thickness pertaining to the "ohms per square" of a sheet.

The resistance between A′-B′ will be the same as that between A-B, since, although the length has been doubled (which alone would double the resistance), the cross-sectional area has also been doubled (which alone would cut the resistance in half).

Thus, *for a given thickness* we can refer to the "ohms per square" of a sheet. Typical values run from 50 to 250 ohms per square, depending upon the thickness for a P-type diffused material, and from 50 to 10 K ohms per square for thin films.

Thin films are only a few microns thick (using such materials as Nichrome - 400 ohms per square), whereas thick films are several microns thick (resistances up to 20,000 ohms per square are available). Both are less temperature-sensitive than the diffused resistors. Thick film resistors may be trimmed to provide values within ±1% of the desired value.

The area required for a large resistance may become prohibitive if a norma film resistor is considered. (In the tion example above it required 50 sq. provide a resistance of 10 K ohms using 20 ohms/sq. of diffused material). One way around this is to use a "pinch resistor." As shown by Fig. 24-12, this consists of a normal P-type diffusion with an N^+-type layer on top. This reduces the effective cross-sectional area of the resistor by eliminating the conductivity paths near the surface of the P-type diffusion.

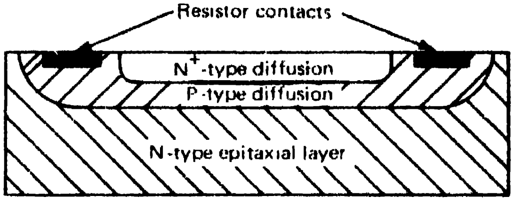


Fig. 24-12. Cross-sectional view of a pinch resistor.

In this way, practical values of resistors from 10 K to 500 K ohms can be obtained due to sheet resistivities in the order of 5000 ohms per square. Thus a 10 K ohm pinch resistor would occupy approximately one-twentieth the space required for the P-type diffused resistor. However, the penalty paid for this saving of space is a very poor tolerance on nominal resistance values, as seen in Fig. 24-13.

The wide tolerances in resistance values shown arise from variations in the diffusion and etching processes. However, the *ratio* of two integrated-circuit resistors of the same type and constructed at the same time on the same chip can be closely controlled—typically within 2 to 3 percent in P-type diffused resistors. Since this is also true of the characteristics of transistors, most IC designs rely upon resistance ratios and *symmetrical* circuits, as in the differential amplifier circuit in the previous chapter.

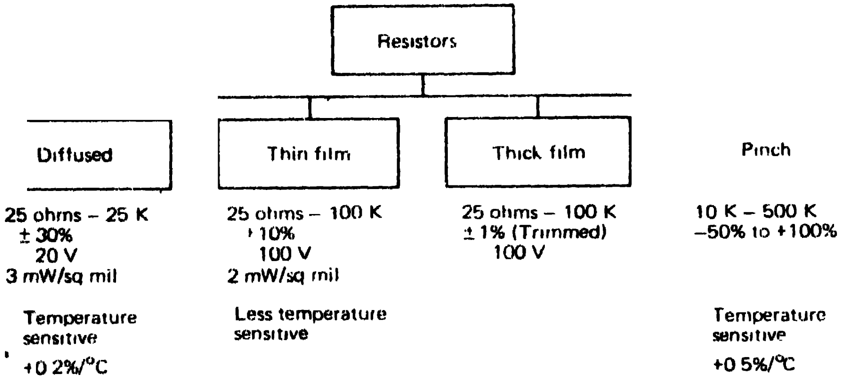

Resistors
Diffused
Thin film
Thick film
Pinch
25 ohms – 25 K
± 30%
20 V
3 mW/sq mil
Temperature sensitive
+0.2%/°C
25 ohms – 100 K
± 10%
100 V
2 mW/sq mil
Less temperature sensitive
25 ohms – 100 K
± 1% (Trimmed)
100 V
10 K – 500 K
–50% to +100%
Temperature sensitive
+0.5%/°C

Fig. 24-13. Showing the types of IC resistors available and some of their characteristics.

Capacitors

As shown by Fig. 24-14, capacitors may be obtained in essentially two different ways.

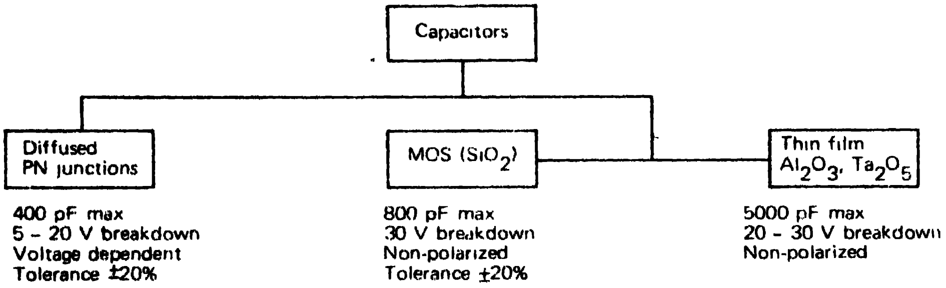

Capacitors
Diffused PN junctions
MOS (SiO2)
Thin film Al2O3, Ta2O5
400 pF max
5 – 20 V breakdown
Voltage dependent
Tolerance ±20%
800 pF max
30 V breakdown
Non-polarized
Tolerance ±20%
5000 pF max
20 – 30 V breakdown
Non-polarized

Fig. 24-14. Showing the types of IC capacitors available and some of their properties.

As explained in Chapter 5, a reverse-biased PN junction provides a capacitive effect. This junction may be provided by a transistor in which typically 0.4 pF per square mil exists for an emitter-base junction and 0.1 pF per square mil for a base-collector junction.

There is only a subtle difference between the the MOS capacitor and the true thin-film capacitors. The MOS employs a low-resistivity N-type diffused layer as one plate of the capacitor (upon which a silicon dioxide layer and an aluminum plate are added); in the thin-film capacitor, a separate layer of aluminum is first provided, on which the aluminum or tantalum oxide is deposited, followed by a second aluminum layer. Examples of these are shown in the next section.

The area required for high values of capacitance soon becomes uneconomically large, so that circuit design is modified to accept values in the 3 to 30 pF range. Where this is not possible, it is more economical to connect discrete components exterior to the IC chip.

Some idea of the relative area used by diffused resistors, bipolar transistors, and PN junction capacitors is provided by the following:

COMPONENT	RELATIVE AREA
Transistor	1
1000-ohm resistor	2
10-pF capacitor	3

Since area determines cost, active devices are used wherever possible, to reduce the number of large-area passive components.

Inductors

To date, no satisfactory method of fabricating appreciable values of inductance or transformers has been found, so that these also are usually found exterior to the IC where their use cannot be designed out of the circuit.

Diodes

Diodes are obtained by using the PN junctions available in a diffused transistor structure. See Fig. 24-15. In (a), with the base and collector joined together, the base-emitter junction when reverse biased is often used to provide the equivalent of a Zener diode, with $V_Z = 7$ V and a dynamic impedance between 60 and 100 ohms. Although this configuration may also be used as a diode in the forward direction, larger breakdown voltages are provided by the collector-base junction in (b), although at a higher forward voltage drop.

Field-Effect Transistors

As was noted in a previous chapter, MOSFETs are a product of integrated circuit technology. (Review the diagrams of the MOSFETs in Chapter 22.) They actually have some distinct advantages over the bipolar transistor from a fabrication standpoint, requiring less than one-tenth of the area and requiring no isolation from adjoining components, because of the reverse-biased PN junction they utilize in normal operation. Their fabrication will be referred to again in Section 24-5.

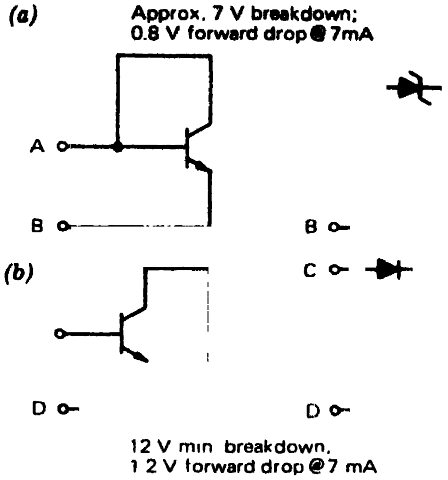


Fig. 24-15. (a) A transistor with collector shorted to base to provide a diode with a Zener reverse breakdown effect. (b) The transistor's collector-base junction provides a diode with a higher breakdown voltage.

Bipolar Transistors

Both NPN and PNP transistors may be fabricated in ICs by diffusion, but the NPN is often used when isolation between components is achieved by reverse biasing the substrate (P-type) and the collector (N-type) junction. This is done by connecting the P-type substrate to the most negative point in the circuit. This is shown in the following section.

24-4. METHODS OF FABRICATING INTEGRATED CIRCUITS

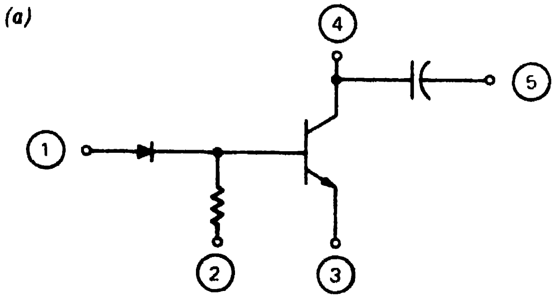
In order to illustrate the components mentioned in the previous section, let us briefly look at the different methods of fabricating integrated circuits.

Monolithic Integrated Circuit

Literally, monolithic means "one-stone," from which the whole circuit is manufactured using *only* diffusion methods and no "add-ons." The circuit "fabricated" in Section 24-2 was an example of a monolithic circuit. Another example is shown in Fig. 24-16 and includes a few more components. To avoid confusion, the metallization between areas and bonding pads has not been shown. The emphasis is on component structure.

Note the method of isolation shown in Fig. 24-16. By connecting the P-type substrate to the most negative point in the circuit, the PN junctions separating the components are reverse biased, providing dc isolation in excess of 10 M ohms. Observe that a parasitic PNP transistor is formed by the P-type substrate and the adjoining layers of N- and P-type materials. Thus, in normal circuit operation, the base-collector junction of the NPN transistor should not be forward biased in excess of 0.6 V; otherwise the parasitic transistor will be turned on, and high current will flow into the substrate.

Another method of providing isolation between components involves the use of a dielectric such as silicon dioxide to surround the components in the substrate. No reverse bias between substrate and circuit elements is then necessary. And this method also makes it a relatively simple matter to fabricate both NPN and PNP type transistors on the same silicon substrate.



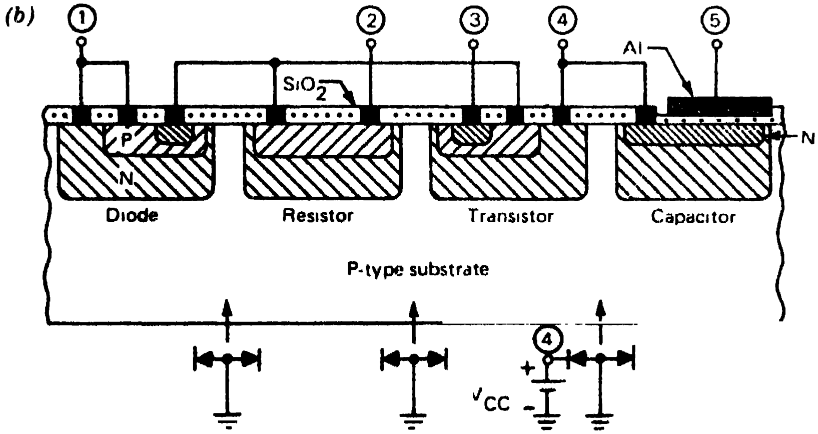


Fig. 24-16. (a) Circuit to be built using the monolithic approach. (b) Cross-section of the monolithic IC showing how the N-type islands resting in the P-type substrate provide PN junctions, which, when reverse-biased, effect isolation of one island or component from the other.

Hybrid Integrated Circuits

Thin-Film Integrated Circuit

Since no *satisfactory* thin-film active devices have yet been developed, transistors must be added to thin-film integrated circuits just as if they were discrete components. This is shown in Fig. 24-17.

Note that a passive substrate has been used here, and a diffused transistor whose collector is the substrate provides a bottom connection to the resistor.

Compatible Silicon-Thin-Film Integrated Circuit

This circuit consists of thin films (to provide resistors and capacitors) added to the completed diffused portion of the circuit, as illustrated in Fig. 24-18.

The process combines the best features of both the monolithic and the thin-film.

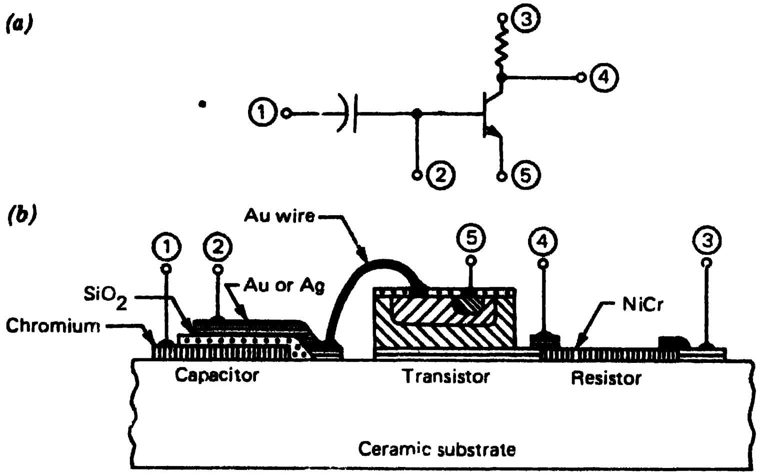

(a)
③
④
①
②
⑤
(b)
Au wire
①
②
⑤
④
③
Au or Ag
SiO₂
Chromium
NiCr
Capacitor
Transistor
Resistor
Ceramic substrate

Fig. 24-17. (a) Circuit to be implemented. (b) Thin-film hybrid integrated circuit. Note the "discrete" transistor (Courtesy of Integrated Circuit Engineering Corp.)

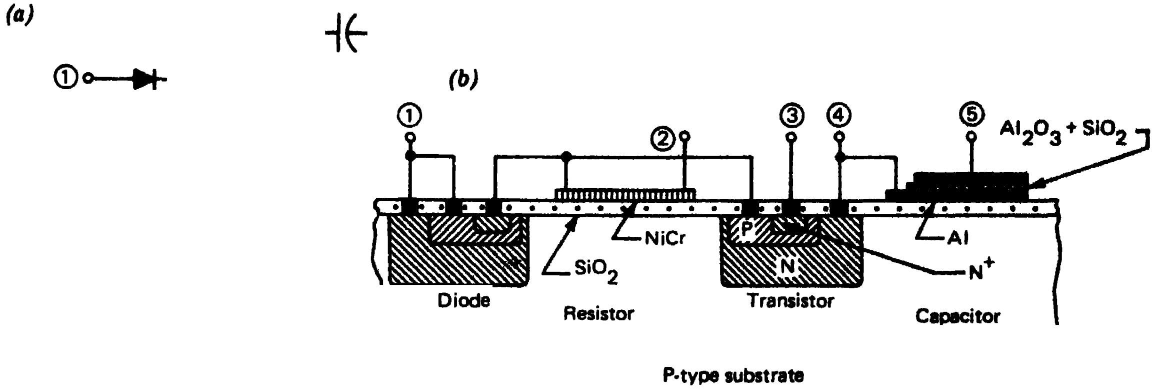

(a)
①
(b)
①
②
③
④
⑤
Al₂O₃ + SiO₂
P
N
NiCr
SiO₂
Al
N⁺
Diode
Resistor
Transistor
Capacitor
P-type substrate

Fig. 24-18. (a) Circuit to be implemented. (b) Cross-sectional view through a compatible silicon-thin-film hybrid integrated circuit. (Courtesy of Integrated Circuit Engineering Corp.)

Multichip Integrated Circuit

This circuit is made by interconnecting each chip as if it were a discrete component. The method is used mostly for small-run custom-built requirements and interim design for complete monolithic circuits.

24-5. COMPLEMENTARY SYMMETRY MOS INTEGRATED CIRCUIT (COS/MOS)*

All of our discussion thus far may have implied linear circuits were being constructed, but all of the fabrication techniques covered apply equally well to digital devices. Although these are to be covered in detail in the next chapter, a digital device consists essentially of switching circuits in which an active device is either on or off. They are then nonlinear in nature.

These nonlinear devices may include the bipolar type of transistors, along with diodes, resistors, and capacitors, but for many applications the MOS devices provide a significant advantage. Although somewhat restricted in switching speed due to the capacitances involved (compared with corresponding bipolar circuits), MOSFETs require far less space (approximately 5 sq. mils of area compared with 50 sq. mils for a triple diffused bipolar transistor).

An additional advantage of low power dissipation is found in the *co*mplementary symmetry *MOS*FET construction, in which both P-channel and N-channel enhancement-type transistors are diffused into the same chip. They are then interconnected to form a basic inverter circuit, as shown in Fig. 24-19.

Operation of the COS/MOS Inverter Circuit

As can be seen from the diagram, the complementary MOSFETs are connected in series, with their gates connected together and their drains connected together to form input and output terminals, respectively.

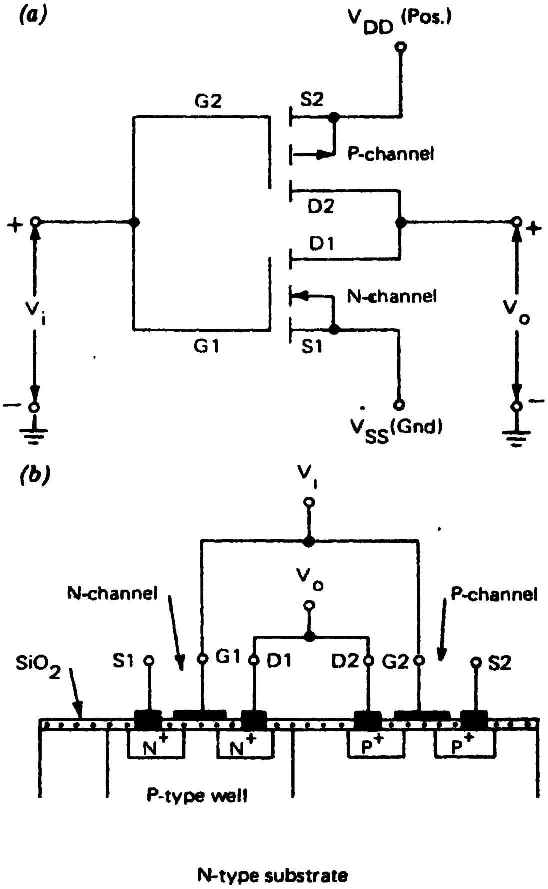


Fig. 24-19. (a) Schematic diagram of a COS/MOS inverter circuit utilizing two series-connected complementary enhancement MOSFETs.
(b) Simplified diagram of the IC structure of the COS/MOS inverter.

The operation of the circuit is as follows:

(1) Consider V_{DD} to be connected to a positive voltage between 5 and 15 volts and V_{SS}

---

*This is a trade mark used by RCA who pioneered this type of construction. Motorola refers to their complementary MOS circuits as MCMOS.

to be connected to ground.

(2) If $V_i = 0$ V, the gate-to-source voltage of the P channel equals the supply voltage. That is, $V_{G2S2} = -V_{DD}$, with the result that the P channel is on, providing a low resistance path between V_{DD} and the output. The N-channel transistor is off, since $V_{G1S1} = 0$ V, and a high resistance is offered between the output and ground. Thus $V_o = V_{DD}$.

(3) If V_i is made positive and equal to V_{DD}, the P channel is switched off, since $V_{G2S2} = 0$ V, and the N channel switched on, since $V_{G1S1} = V_{DD}$. The low resistance path of the N channel makes V_o approach 0 V. Thus, an inverting action results similar to that in Section 18-9.

Characteristics of the COS/MOS Inverter Circuit

The manner in which the output voltage varies with the input voltage is shown in Fig. 24-20 for several supply voltage values and temperatures. These curves indicate the effect of the threshold voltage that is characteristic of enhancement MOSFETs, as discussed in Section 22-2. Switching from one state to the other takes place typically at 45 percent of the supply voltage, providing a high noise immunity (immunity to false switching arising from transient voltage spikes, etc.).

The fact that one transistor is always in the off state means that the quiescent dissipation of the inverter (no load connected and no change in input) is extremely small. It is actually equal to the product of the supply voltage and leakage current, typically 1 microwatt at 10 volts. This is the outstanding advantage of the COS/MOS IC, its dissipation being far less than that of comparable single P-channel MOS and bipolar ICs.

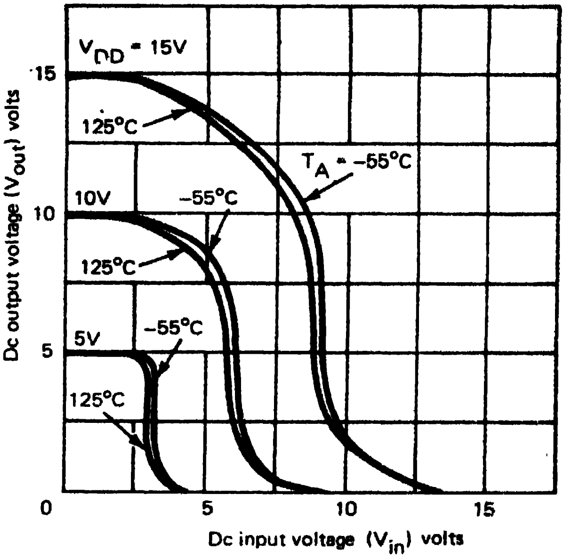


Fig. 24-20. Showing transfer characteristic curves of a COS/MOS inverter circuit.

Its dissipation under load depends upon the rate of switching and the amount of capacitive load involved. That is, if this inverter (or any more involved circuitry built from a number of differently interconnected inverters to perform other functions) is used to drive other COS/MOS ICs, the load is predominantly capacitive due to the insulated gate inputs. Consequently, COS/MOS ICs have high "fanout" capability (the number of units that can be connected to its output), but the rise and fall times (in the vicinity of 100 nsec) will increase with load capacitance.

Fabrication of the COS/MOS Inverter Circuit

It can be seen from Fig. 24-19(b) that the starting point in the fabrication sequence is an N-type substrate. Into this is diffused a P-type well, to provide the effective substrate for the N-channel transistor. Further alternate diffusions provide the drain and source for each device, although at the same time additional areas (not shown) are diffused to provide the gate protection diodes as covered in Section 22-6.

It should be noted that the silicon dioxide that was used solely for the diffusion masking process in the bipolar transistors is now also used as an integral part of the MOSFET gate construction.

Although the diagram in Fig. 24-19(b) indicates interconnection between the complementary devices as occurring between "side-by-side" units, this is not so. The transistors are laid out in checker-board fashion so that easy interconnection by straight line metallization paths is possible from one transistor to its complementary counterpart directly "in front of" or "behind" it.

It is interesting to note that the N-channel device exhibits a higher transconductance (due to the greater mobility of electrons) than is found in the P-channel type. Thus for matched, balanced conditions the P-channel MOSFET must occupy approximately twice the channel width found in the N-channel unit for the same length.

24-6. LARGE-SCALE INTEGRATION (LSI)

Much of the reliability credited to ICs arises from the fact that interconnections are made within the chip and are not subject to the problems arising from discrete connections such as vibration, fatigue, cold-solder joints, etc. Thus it was only natural to build on the same chip many interconnected circuits to provide some overall system or subsystem. This has been especially effective in the digital area and has been enhanced by the use of MOS devices instead of bipolar transistors (approximately one-tenth the area).

For example, the basic inverter circuit referred to in the previous section may be used to make any of the basic *logic gates* used in digital computers and many control functions. These will be discussed in detail in the next chapter, but basically a logic gate is used in decision-making circuits to control the path of signals and provide some output in accordance with some required sequence of inputs. Loosely speaking, an inverter is sometimes referred to as a *gate* but this term is generally applied to circuits with more than one input.

Large-scale integration (LSI) refers to applications where more than 100 gates are fabricated on the same chip. If the number of gates is more than 12 but less than 100, this is referred to as *medium-scale integration* (MSI). Note that the number of components may be much higher than the number of gates. For example, 100 inverters consist of 200 components. Or, 100 NOR gates (see Chapter 25), each of which may be obtained by interconnecting four COS/MOS inverters, would consist of 400 components.

The closest thing to a complete computer on a chip was announced in early 1972, when the central processor of a computer, containing 2800 bipolar transistors, was fabricated on a single LSI chip.

A more recent application involves MOS/LSI technology in a pocket electronic calculator. See Fig. 24-21. Each of the LSI circuits, presently believed to be the largest in volume production in the world, contains the equivalent of 6,000 transistors, a total of 30,000. See Fig. 24-22 and Table 24-1. This is an example where both the increased reliability and the reduced size provided by ICs are used to good advantage.

24-7. SUMMARY

(1) Integrated circuits compose the major portion of the field of microelectronics and may consist of film, monolithic, or hybrid circuits.

(2) A monolithic IC consists of active and passive components formed by diffusions into a single silicon chip, with interconnections provided by an aluminum metallization process.

(3) The monolithic fabrication process consists of wafer preparation, epitaxial growth,

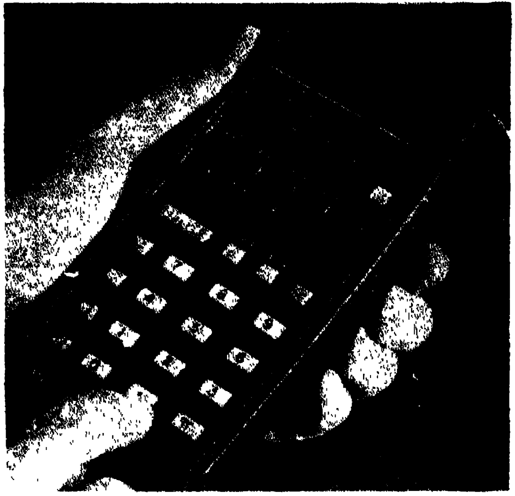


Fig. 24-21. POCKET ANSWER MACHINE: The HP-35, a tiny (3 in. X 6 in. X 1 in.) electronic calculator from Hewlett-Packard, weighs only nine ounces and fits into a shirt pocket. Battery-powered unit performs all trigonometric and logarithmic functions, square root, arithmetic functions, and several other mathematical operations, each with a single key-stroke. It can be likened to a fast, extremely accurate electronic slide rule with solid-state memory similar to those used in computers. Answers appear automatically on the calculator's bright red light-emitting-diode display, which can show numbers having up to 10 digits plus two-digit exponents and appropriate signs. (Courtesy of Hewlett-Packard Co.)

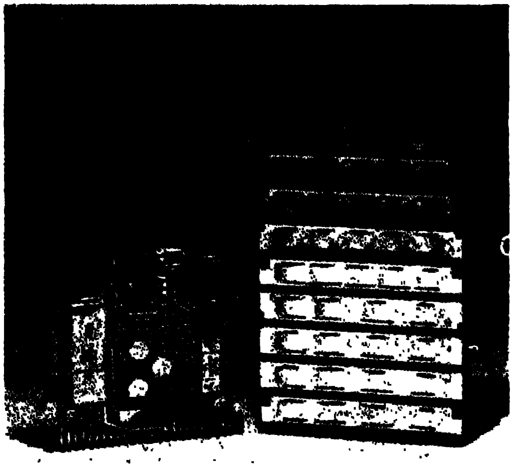


Fig. 24-22. View of the two circuit boards of the HP-35 electronic calculator.

On the circuit board on the left can be seen three 10-pin TO-5 ICs that constitute the Instruction ROM (Read Only Memory) chips. On the right of the circuit board is the 16-pin DIP making up the Arithmetic and Register chip, while to the left is the 28-pin DIP for the Timing and Control chip.

At the top of the other circuit board on the right can be seen the back of the three 14-pin DIP 5-digit light-emitting-diode (LED) clusters. Beneath them are the decoder/driver packages. The rest of the board is filled with the rear contacts for the unit's 39 keys. (Courtesy of Hewlett-Packard Co.)

TABLE 24-1 DETAILS OF THE LSI CHIPS USED IN THE HP-35 ELECTRONIC CALCULATOR (COURTESY OF HEWLETT-PACKARD CO.)

CIRCUIT TYPE	LOGIC GATES	FLIP-FLOPS	ROM STORAGE	PACKAGE	FINAL CHIP SIZE (IN MILLI-INCHES)	NOMINAL POWER (IN MILLI-WATTS)
Instruction ROM	100	30	256 X 10	10-pin TO-5	126 X 134	3
Arithmetic and Register	335	460	–	16-pin C-DIP	156 X 200	20
Timing and Control	340	55	25 X 58	28-pin C-DIP	150 X 212	50

diffused isolation, base and emitter diffusions, pre-ohmic etch, metallization, circuit probing, dicing, mounting and packaging, wire bonding, encapsulation, and final testing.

(4) Each diffusion process involves the following items: silicon dioxide layer, photoresist, mask, ultraviolet exposure, etching, scrubbing, and diffusion.

(5) Practical IC resistance values obtainable range from 25 ohms to 500 K ohms, depending upon the sheet resistivities measured in ohms per square. P-type diffused resistors have values from 50 to 250 ohms per square and pinch resistors have a value of 5000 ohms per square.

(6) Because the tolerances of resistance values are at best ±30%, ICs are designed to utilize resistance *ratios*, which may be controlled to within 3 percent.

(7) Capacitors may be obtained by utilizing the PN junctions in transistor type structures or the MOS capacitive effects employing the silicon-dioxide layer. Practical values range from 3 to 30 pF because of the excessive area used for larger values.

(8) Diodes are also obtained by using the PN junctions available in a diffused transistor structure.

(9) Integrated circuits may be of the monolithic or hybrid type, with the latter being a mixture of thin-film and diffused components.

(10) Complementary symmetry MOS ICs (COS/MOS) involve both a P-channel and an N-channel MOSFET fabricated on the same chip.

(11) By suitably interconnecting COS/MOS ICs, an inverter circuit with very low quiescent dissipation can be made available, and by further interconnections of many inverters many of the logic gates found in digital systems may be implemented.

(12) Large scale integration (LSI) involves the fabrication of 100 or more logic gates on a single chip, while medium scale integration (MSI) is defined as more than 12 but less than 100 gates on a single chip.

QUESTIONS

(Many of the following multiple choice questions are courtesy of Texas Instruments Incorporated.)

Q24-1. The term microelectronics is restricted to:

A. Miniature discrete-component circuits.
B. Monolithic ICs.
C. Circuits made by thin-film, thick-film, or semiconductor techniques.

Q24-2. Monolithic IC wafers are typically:

A. $1\frac{1}{2}$ in. diameter.
B. 50 mils square.
C. 5 mils square.

Q24-3. N-type silicon is:

A. Formed by doping pure silicon with phosphorus.
B. Formed by doping pure silicon with boron.
C. Nonconducting.

Q24-4. In a silicon monolithic IC, isolation is:

A. Unnecessary because the silicon substrate has a high resistivity.
B. Necessary because the silicon is electrically conducting.
C. Obtained by mounting separate chips on an insulating base.

Q24-5. Silicon dioxide is used as a:

A. Diffusing element.
B. Mask against diffusion.
C. Contact material.

Q24-6. The photoresist process:

A. Prevents photoresponse.
B. Controls the etching of silicon oxide from selected areas.
C. Deposits a material which prevents diffusion in selected areas.

Q24-7. Photomasking:

A. Controls the depth of diffusion.
B. Is used to remove selected regions of silicon oxide.
C. Forms an insulating layer that prevents diffusion in selected areas.

Q24-8. The etching process involves:

A. Removal of the photoresist.
B. Removal of the silicon dioxide coating.
C. Scribing lines on the wafer prior to dicing.

Q24-9. Pre-ohmic etch:

A. Provides openings to make metal contact with the diffused areas.
B. Removes excess resistance from diffused resistors.
C. Involves the trimming of thick film resistors to ±1%.

Q24-10. Metallization is used to provide:

A. Solely bonding pads.
B. Connection from the base to the bonding pad.
C. Connections from the diffused areas up to and including the bonding pads.

Q24-11. Circuit probing is:

A. First done after dicing.
B. Done on the complete wafer to mark faulty circuits.
C. A destructive test to remove faulty circuits.

Q24-12. The IC chip is mounted to its package base by:

A. Thermocompression bonding.
B. Ball bonding.
C. Alloying or cementing.

Q24-13. Wire bonding involves:

A. Thermal compression to connect the IC pads and the package leads.

B. The cementing of the IC die to the header.
C. The placement of the leads through the header.

Q24-14. Functional testing of the IC wafer is performed during:

A. Probe testing.
B. Final testing.
C. Both A and B.

Q24-15. Low yields of acceptable ICs are:

A. Inevitable.
B. Due to an accumulation of process errors, crystal imperfections, and contamination.
C. Mainly due to losses at final test.

Q24-16. Diffused IC resistors:

A. Need no isolation.
B. Can be produced with an accurate ratio between resistance values.
C. Can be accurately reproduced in absolute resistance values.

Q24-17. Thin-film circuits are:

A. Formed with evaporated films less than 0.1 mil thick.
B. Less than an eighth-inch square.
C. Printed using metallized ink.

Q24-18. IC capacitors:

A. Can be made by using silicon oxide as the dielectric.
B. Cannot be made with diffusion techniques.
C. Can only be made with very high capacitance values.

Q24-19. The MOS transistor:

A. Has only one PN junction.
B. Has a gate electrode in direct contact with the silicon.
C. Conducts when sufficient voltage is applied to the gate electrode.

Q24-20. The advantage of a MOSFET over a bipolar transistor is:

A. Increased operating speed.
B. Smaller fabrication area for diffusion.
C. Increased reliability.

Q24-21. Transistors for monolithic ICs:

A. Are identical with discrete planar transistors.
B. Use the isolation junction as the collector junction.
C. Are similar to discrete planar transistors, but have the collector contact on the top surface.

Q24-22. In a monolithic IC:

A. All circuit elements are formed in a single wafer of semiconductor.
B. Active elements are assembled as separate wafers.
C. Circuit elements are assembled on ceramic substrates.

Q24-23. A multichip circuit is:

A. A monolithic silicon wafer with thin-film components.
B. Several interconnected monolithic circuit chips.
C. A circuit with each circuit element on a separate chip.

Q24-24. An integrated circuit contains 30 gates, each of which consists of 4 components. This is an example of:

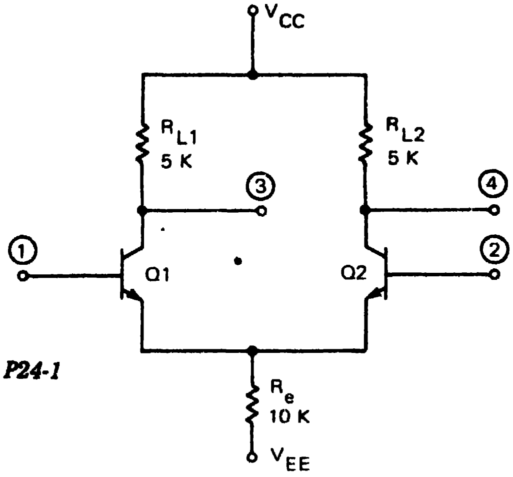
A. Large scale integration (LSI).
B. Medium scale integration (MSI).
C. Small scale integration (SSI).

Q24-25. The primary advantage of integrated circuits over discrete components is:

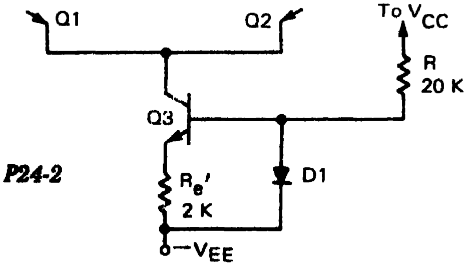
A. Reduced size.
B. Increased reliability.
C. Increased operating speed.

PROBLEMS

P24-1. Design a monolithic IC to implement the following schematic of a simple differential amplifier:



P24-1



P24-2

Show suitable cross-sections to illustrate the diffused areas and a plan view to indicate the metallization interconnections as shown in Fig. 24-8. Provide bonding pads for the voltages and the terminals indicated. Suggest sheet resistivities for the resistors.

P24-2. Repeat P24-1 assuming R_e is replaced by the transistor-diode combination shown.

P24-3. Describe the operation of a COS/MOS inverter circuit in which the N-channel device has its source connected to V_{DD} and the P-channel device's source is connected to ground. Specify the polarity of V_{DD} and of the input and output voltages.

P24-4. (a) If the fabrication of an IC requires 35 steps each of which has a 90% accuracy, what is the total yield?

(b) If it is required to increase the yield in (a) to 25%, what must be the average accuracy of each step?

(c) If the accuracy in (b) is unrealistically high, what is the maximum number of steps the process must be reduced to in order to achieve a total yield of 25% at a 94% accuracy?

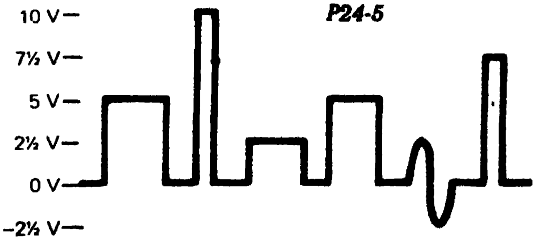
P24-5. Consider the inverter circuit in Fig. 24-19(a). Assume V_{DD} = +5 V and the following train of pulses is applied to the input.

(a) Draw the corresponding output from the inverter.

(b) How will the output waveform be modified if V_{DD} is increased to 15 V? Use the curves in Fig. 24-20.

P24-6. Assume the pulse train in P24-5 is negative going instead of positive going. Draw the output waveform from the inverter in P24-3, with V_{DD} = -5 V.

P24-5



25

NONLINEAR INTEGRATED CIRCUITS I: COMBINATIONAL DIGITAL CIRCUITS

25-0. INTRODUCTION

The first major application of integrated-circuit technology was in the field of digital devices, where diodes and transistors were used to perform switching functions in what are called logic gates. These form the heart of digital computers and are widely used in control systems. Basically, a gate is a circuit to provide an output which is dependent upon the manner in which the inputs are applied.

As may be seen from Fig. 25-1, logic gates may be subdivided into the combinational and sequential types. In the former, the output at a given instant depends only upon the inputs at the same instant, the previous history of the inputs to the gate is immaterial. In the sequential gates, the output depends upon the order or sequence in which inputs are applied, and sequential gates are thus said to possess a *memory* function.

In this chapter, we shall examine the manner in which the combinational gates are implemented and the different IC logic *families* available. Chapter 26 covers the sequential gates.

25-1. THE BINARY CONCEPT

In digital systems, there are two possible states or conditions. For example, a switch may be closed or open, a situation true or false, a voltage level high or low, etc. The two states may be referred to as 1 and 0, respectively. That is, a closed switch or the existence of a signal may be referred to as a binary 1, and the open switch or the absence of a signal as a binary 0.

In decimal arithmetic, ten numerals are required, 0,1,2,...,9 and the number 1973 is a representation of:

$$1 \times 10^3 + 9 \times 10^2 + 7 \times 10^1 + 3 \times 10^0,$$

with 10 being the arithmetic base.

In binary arithmetic only two numerals are required, 0 and 1, since 2 is the base. Thus 1001101 has the following meaning:

$$1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 1 \times 2^2$$

$$+ 0 \times 2^1 + 1 \times 2^0$$

$$= 64 + 0 + 0 + 8 + 4 + 0 + 1$$

$$= 77,$$

and is equivalent to 77 in decimal notation.

Note that the most significant digit (or bit in binary) MSB, is on the left, as in the decimal, and the least significant bit, LSB, is on the right. Conversion from decimal to binary is achieved by repeated division of the decimal number by 2 and writing all the remainders in sequence with the last remainder the MSB. (Try it for 77.)

In a digital system, voltages representing analog quantities such as speed, pressure, temperature, etc., must be converted to binary (or some type of binary code). Such a process is called analog-to-digital conversion (A/D). After the signals have been processed or computation has taken place, the reverse process of digital-to-analog conversion, D/A, is often necessary to display the information in readable form. These conversions are quite complex and will not be discussed here at all. (A binary coded decimal (BCD) to decimal readout is mentioned in Chapter 26.)

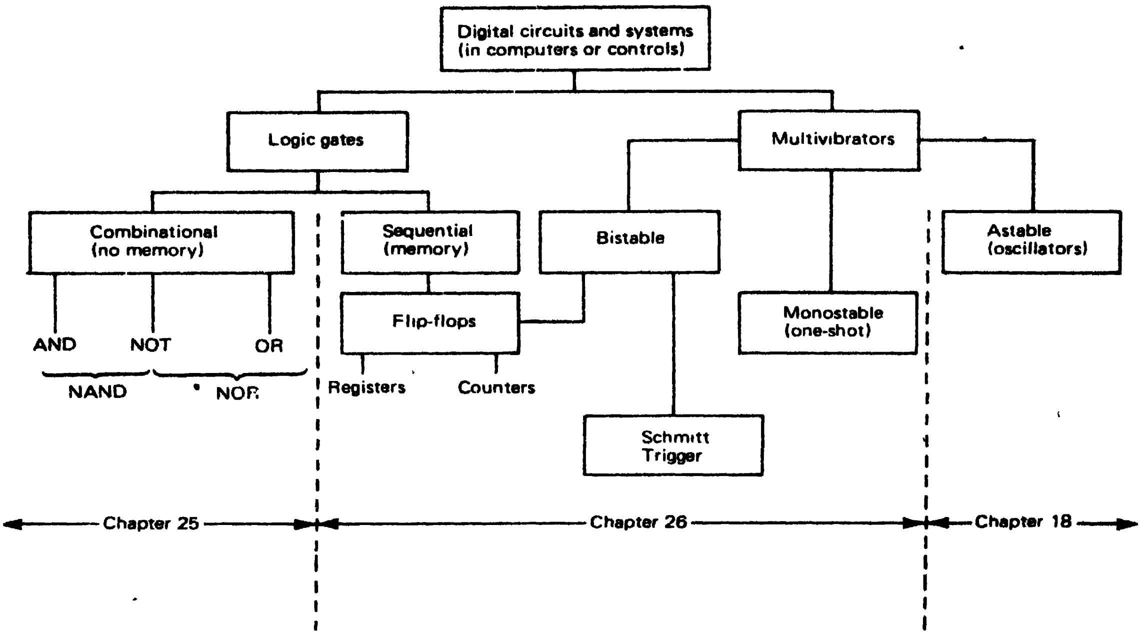

Digital circuits and systems (in computers or controls)
Logic gates
Multivibrators
Combinational (no memory)
Sequential (memory)
Bistable
Astable (oscillators)
Flip-flops
Monostable (one-shot)
AND
NOT
OR
NAND
NOR
Registers
Counters
Schmitt Trigger
Chapter 25
Chapter 26
Chapter 18

Fig. 25-1. Digital circuit "tree."

Logic Systems

A digital bit of information may be represented by a voltage *level* in a dc system or a voltage *pulse* in a dynamic system.

When voltage levels are used, two methods are possible, positive logic and negative logic. In either method, two different voltage levels are used to represent the 0 and 1. In positive logic, the more positive of the two voltage levels represents the 1, while in negative logic the more negative voltage represents the 1. In neither case does the 0 state *have* to be represented by 0 volts, although in some systems the two may coincide. For example, one system uses 0 to 1 V as the lower level or 0 and 3 to 5 V as the higher level or 1. In this way, the logic levels are clearly distinguishable. Another system uses -1.55 V as the typical value for 0 and -0.75 V for 1. Note that both of these are examples of positive logic. To avoid confusion we shall stay with the positive logic, although the type of logic used often determines the exact function provided by a circuit.

In a positive-logic dynamic system using pulses, a dynamic 0 is a negative-going voltage excursion and a dynamic 1 is a positive-going voltage excursion. See Fig. 25-2.

When pulses are used for the binary information, it is often required that they have very fast rising and falling edges. Circuits to provide this are considered later.

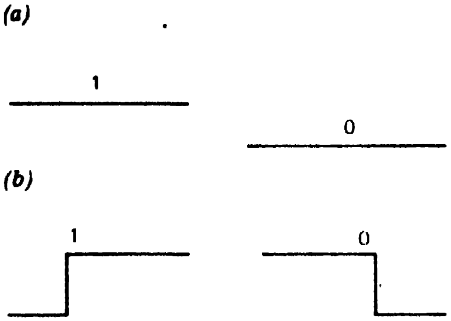
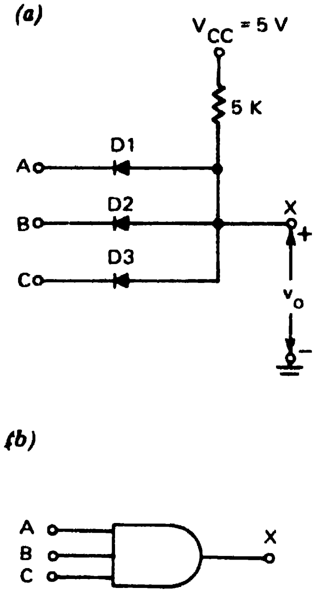


Fig. 25-2. Diagrams showing a binary 1 and 0 in (a) voltage level system and (b) voltage pulse system.



(c)

Inputs			Output
A	B	C	X
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Fig. 25-3. (a) Schematic circuit of a positive logic diode AND gate. (b) Standard symbol for an AND gate. (c) Truth table for a three-input AND gate.

25-2. BASIC LOGIC GATES

The AND Gate

The AND gate is a circuit to provide an output (a logical 1) when all the inputs are simultaneously present (all have logical states of 1). A diode implementation of a positive logic three-input AND gate is shown in Fig. 25-3 along with its symbol and *truth table.* This table shows the output for all the possible combinations at the input. Note that 0 = 0 V and 1 = +5 V.

The AND gate may be thought of as being electrically equivalent to a circuit which has three switches A, B and C connected in *series.* Only if A *and* B *and* C are closed will an output result.

In Fig. 25-3, X = 1 if A = B = C = 1. The logical equation relating X to A, B and C may be given using Boolean algebra (developed by George Boole in the middle nineteenth century). In this system the dot (·) or cross (×) when placed between input symbols indicates the AND function:

$$X = A \cdot B \cdot C = A \times B \times C = ABC. \qquad (25\text{-}1)$$

Thus, the value of X is 1 (present) when A *and* B *and* C are 1 (present), as indicated by:

$$1 = 1 \cdot 1 \cdot 1. \qquad (25\text{-}2)$$

But if A = 1, B = 1, and C = 0, then,

$$1 \cdot 1 \cdot 0 = 0, \qquad (25\text{-}3)$$

and no output results as indicated by the truth table. (Note that all the possible combinations of the inputs in the truth table are obtained by counting binarily from 0 through 7).

The operation of the diode AND circuit in Fig. 25-3(a) is as follows:

(a) If A = B = C = 0 = 0 V, all three diodes are forward biased, and $v_o \approx 0.6$ V so that X = 0.
(b) If A = B = 1 = 5 V (or more), diodes D1 and D2 do not conduct, but, if C = 0 = 0 V, D3 will still be forward biased. Thus, v_o will remain at approximately 0.6 V and X = 0.
(c) If A = B = C = 1 = 5 V (or more), all the diodes are reverse biased, so that v_o will rise toward +5 V, and X = 1.

Hence, an output results only when A and B and C are present. Note that the absence of a signal at the input requires that the input be grounded (0); it must not be left floating in this type of circuit (diode logic, DL).

The actual output voltage from the gate will depend upon the load connected to it. The number of gates that a unit can drive, called the *fan-out*, varies with the type of circuit, and may be as low as 5 or as high as 50.

The OR Gate

The OR provides an output when one of the inputs is in the 1 state. A two-input OR gate is shown in Fig. 25-4 with its symbol and truth table. Although diodes have been used in this circuit, we shall see later that the OR function is more readily obtained using transistors.

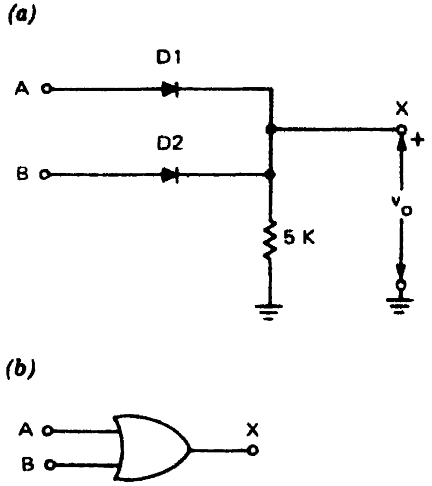
Since an output results from an input at A *or* B (or both), the OR gate may be likened to an electrical circuit in which switches are connected in *parallel* with each other.

The Boolean equation relating the variables in an OR gate is indicated by means of a plus (+) sign between the inputs:

$$X = A + B. \tag{25-4}$$

Thus, if A = 1 and B = 0, then

$$X = 1 + 0 = 1. \tag{25-5}$$



(c)

Inputs		Output
A	B	X
0	0	0
0	1	1
1	0	1
1	1	1

Fig. 25-4. (a) Schematic circuit of a positive logic diode OR gate. (b) Standard symbol for an OR gate. (c) Truth table for a two-input OR gate.

But note that, if A = B = 1, then

$$X = 1 + 1 = 1. \tag{25-6}$$

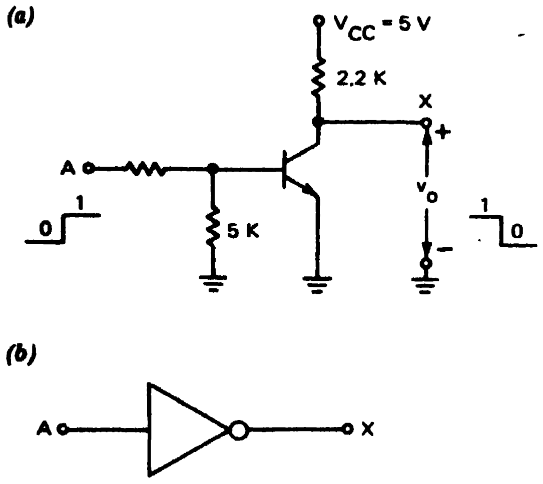
so that the plus sign must not be thought of as an addition process. It represents the logical OR.

The operation of the circuit should be clear. If A = B = 0 (0 V), then v_o = 0 V and X = 0. Observe that the inputs may be left floating if an input is the logical 0. It is not necessary to make sure that a 0 input is grounded, as in the AND gate. If A = 1 = 5 V and B = 0 = 0 V, then D1 will conduct and $v_o \approx 4.4$ V, so that X = 1.

The NOT Circuit

We have already met the NOT circuit in Chapter 18 as an inverter. It is used to *negate* a signal. In Boolean algebra, this is shown by a line drawn over the top of the input. Thus the negation of A is $\overline{A}$, called NOT A or sometimes A-bar.

As you will recall, a transistor driven into saturation is used to achieve this, as shown in Fig. 25-5(a). If A = 0 (0 V or floating), the transistor is held off because $V_{BE} = 0$ V; $v_o = 5$ V and $X = 1 = \overline{A}$. If $A = 1 = 5$ V, the transistor goes into saturation, $v_o \approx 0.2$ V and $X = 0 = \overline{A}$. See the truth table in Fig. 25-5(c).



(c)

Input	Output
A	X
0	1
1	0

Fig. 25-5. (a) Schematic circuit of a positive logic NOT circuit. (b) Standard symbol for an inverter. (c) Truth table for the NOT circuit.

When the inverter is used as a separate entity, its standard symbol is that shown in Fig. 25-5(b). When used in combination with other gates, only the circle portion may be used. An example follows.

Example 25-1. Use the three basic gates, AND, OR, and NOT, to provide a two-input EXCLUSIVE OR gate. That is, a circuit that provides an output when one input or the other is present but not when both inputs are present. (The OR gate in Section 25-2 is sometimes called an INCLUSIVE OR gate since it also provides an output when *both* inputs are available.)

Solution. The first step is to make a truth table See Fig. 25-6(a). Using the truth table, we now construct a Boolean equation relating X to the inputs. It is observed that X = 1 when we have (A AND NOT-B) OR (B AND NOT-A). Alternatively, X = 1 when we have (A AND $\overline{B}$) OR (B AND $\overline{A}$). That is:

$$X = A \cdot \overline{B} + B \cdot \overline{A} = A\overline{B} + B\overline{A}. \qquad (25\text{-}7)$$

A special Boolean expression is used to indicate the EXCLUSIVE OR:

$$X = A \oplus B, \qquad (25\text{-}8)$$

and the symbol is shown in Fig. 25-6(d). Two alternative diagrams for implementing the EXCLUSIVE OR gate are shown in Fig. 25-6 (b), (c).

This circuit, which is sometimes called an inequality comparator or detector, since it only produces an output when the two inputs are different, is a very commonly used circuit in the arithmetic section of a computer. Some manufacturers provide in one 14-lead IC package four such EXCLUSIVE OR gates, called a quad EXCLUSIVE OR gate, such as the Motorola MC771. Each of the four gates consists of the interconnected gates shown in Fig. 25-6(c). Two input leads and one output lead are associated with each gate, and

(a)

Inputs		Output
A	B	X
0	0	0
0	1	1
1	0	1
1	1	0

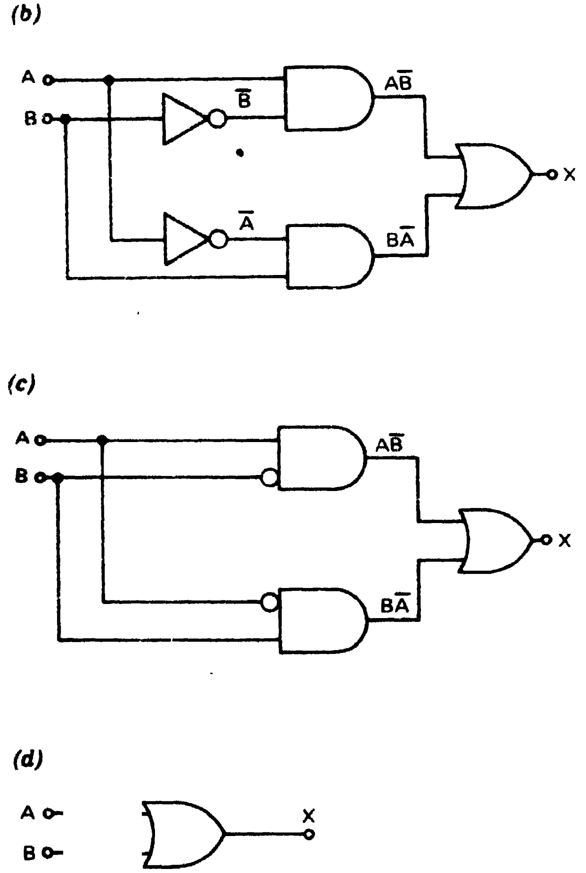


Fig. 25-6. (a) Truth table for EXCLUSIVE OR gate. (b) Implementation using AND-OR-NOT gates. (c) Alternative logic diagram showing inverters as circles. (d) Standard symbol for the EXCLUSIVE OR gate.

provision is made for the connection of V_{CC} and ground, making a total of 14 leads.

The area occupied by the four gates is very small compared with the overall package, but the number of gates provided is limited by the number of connecting leads. Another popular package is the Hex Inverter, such as the Motorola MC789. As the name implies this contains six inverter or NOT circuits.

25-3. DE MORGAN'S LAW

Any logic circuit may be implemented using the basic gates covered. The truth table must be constructed first, followed by writing the Boolean equations. The way in which these equations are written may vary, resulting in different gates being required. In addition, it is often possible to simplify a Boolean equation just as a linear algebraic equation may be simplified, but using different techniques. Although we do not have time to cover this topic fully, one law that is useful in logic building and simplification is De Morgan's Law. Let us illustrate this law by the following example:

Example 25-2. Construct an EXCLUSIVE OR gate using the truth table in Fig. 25-6(a), noting when X is not present.

Solution. We may state that X is NOT present when: (A AND B are present OR when NOT-A AND NOT-B are present). That is:

$$\bar{X} = AB + \bar{A}\bar{B},$$

which is equivalent to:

$$X = \overline{AB + \bar{A}\bar{B}}. \tag{25-9}$$

Note that the negation extends over the whole expression and may be simplified by De Morgan's Law. This states that each input under the negation is negated, AND gates become OR gates, and vice versa. Thus, Eq. 25-9 can be re-written:

$$X = (\bar{A} + \bar{B})(A + B). \qquad (25\text{-}10)$$

Note that a double negation makes no negation. This is implemented in Fig. 25-7(a).

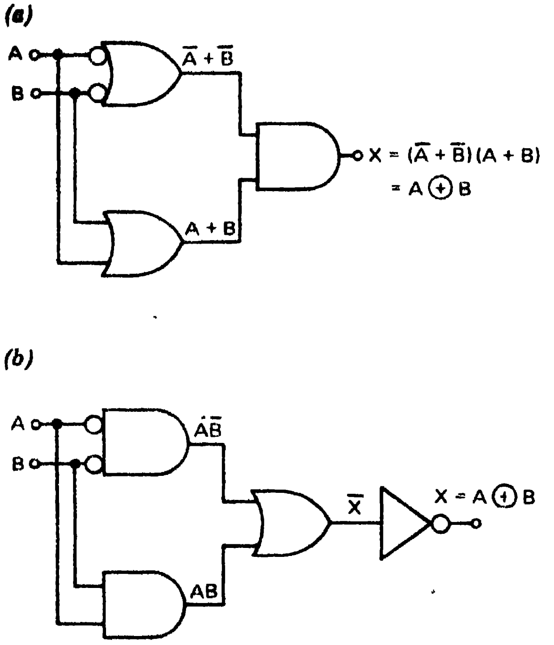


Fig. 25-7. Alternative implementations of an EXCLUSIVE OR gate.

Equation 25-9 may also be modified by negating *both sides* of the equation:

$$\overline{X} = A\ B + A\ B. \qquad (25\text{-}11)$$

If the right hand side of this equation is implemented, the result is NOT–X and must be followed by an inverter to produce the required output, as shown in Fig. 25-7(b).

Although this circuit requires more gates than any of those considered, it has been developed for two reasons. First, it shows that there is more than one way to implement a logic requirement, with some ways being simpler than others. Second, an interesting question is raised in the last two gates of Fig. 25-7(b). Here it is seen that an OR gate is followed by a NOT circuit. Could this NOT circuit be combined or integrated with the OR gate? The answer is yes, the result being a NOT–OR gate or NOR gate. In a similar way a NOT–AND combination is known as a NAND gate. In general these two gates form the majority of the logic gates available. We shall now examine the different methods of implementing them.

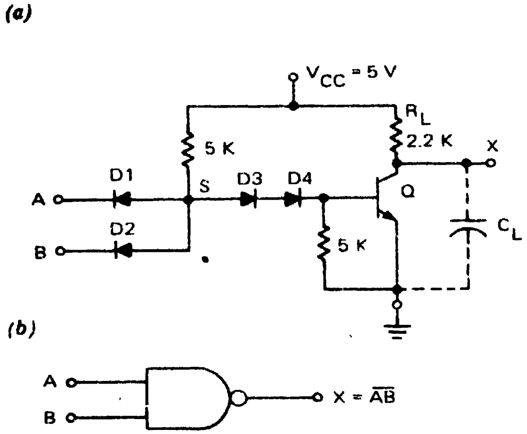
25-4. NAND AND NOR LOGIC GATES

Although logic systems have been developed that use exclusively English (noninverting) logic, such as AND, OR, and NOT, these are usually made from discrete components, including both active and passive devices. IC technology is most efficient in the use of space if active devices are widely used. In addition, active devices provide gain, whereas the signals in the purely diode logic covered earlier tend to degenerate as they pass through a number of gates.

When transistors are used to form logic gates, an inherent inversion often takes place, with the result that NAND or NOR gates result. These may be interconnected to give the AND, OR functions if desired, but generally any Boolean expression written in terms of AND, OR, and NOT may be implemented completely by the use of *either* NAND or NOR gates. This will not be pursued here. Rather, we will look at the different *families* of integrated circuits available and the most common gate that is peculiar to each.

Diode Transistor Logic (DTL) NAND Gate

The NAND gate is most readily constructed when diode transistor logic is used. As suggested by the name, a positive logic DTL NAND gate is a combination of the diode AND gate in Fig. 25-3(a) and the transistor NOT circuit in Fig. 25-5(a), with minor modification as shown in Fig. 25-8. The symbol clearly shows the two-input AND followed by the inverter, and the truth table indicates when outputs occur.



(c)

Inputs		Output
A	B	X
0	0	1
0	1	1
1	0	1
1	1	0

Fig. 25-8. (a) Schematic circuit of a positive logic DTL NAND gate. (b) Standard symbol for a NAND gate. (c) Truth table for a NAND gate.

The truth table may be verified by considering the Boolean expression for the output,

$$X = \overline{AB} = \overline{A} + \overline{B}, \qquad (25\text{-}12)$$

where use has been made of De Morgan's Law. This may be interpreted as saying no output occurs if A and B are present.

Consider that A or B (or both) = 0. Since this will normally be the output from another gate, namely the collector of a transistor, the voltage level representing 0 will be $V_{CE(sat)} \approx 0.2$ V. Say D1 conducts and that its forward voltage drop is 0.6 V. Then $V_S \approx 0.8$ V. However, this voltage is not sufficient to cause diodes D3 and D4 to conduct, since conduction would require $2 \times 0.6\text{ V} = 1.2$ V In addition, it requires approximately 0.5 V cut-in voltage from base to emitter for the transistor to conduct, so that the transistor is in the off state, and X = 1, as required.

Note that the safety margin before the transistor would switch on is $1.2 + 0.5 - 0.8 = 0.9$ V. Thus, a noise immunity of 0.9 V is provided in this condition. That is, a temporary noise spike of less than 0.9 V on one of the inputs would not cause false switching. Had only one diode been used (D4 omitted) the noise immunity would only have been 0.3 V Three diodes would have provided an even greater immunity.

If A = B = 1 (approximately 5 V), diodes D1 and D2 are reverse biased, and the voltage at S is sufficient to drive the transistor into saturation ($V_S \approx 2$ V and $V_{CE(sat)} \approx 0.2$ V). Thus X = 0.

If the circuit of Fig. 25-8(a) is used to drive similar gates, the collector current in the output transistor, Q, will increase (when X = 0) by an amount dependent upon the fan-out, the number of gates driven. This current comes from the forward biased diodes D1, D2, etc., in the driven gates. The fan-out may be increased by the use of an additional transistor, Q1, in place of one of the diodes, as indicated in Fig. 25-9(a). The emitter current of transistor Q1 provides a higher base drive current for the output transistor, Q2, enabling the transistor to handle a higher collector current and drive more gates. The circuit shown in Fig. 25-9(a) is repeated four times on the same chip, to provide a quad 2-input NAND gate in Motorola's MC846, as illustrated by the pin arrangement in Fig. 25-9(b). It is available either in the 14-lead plastic DIP (dual in-line package) or the flat package. Note the indentation and circle depression for package and lead orientation.

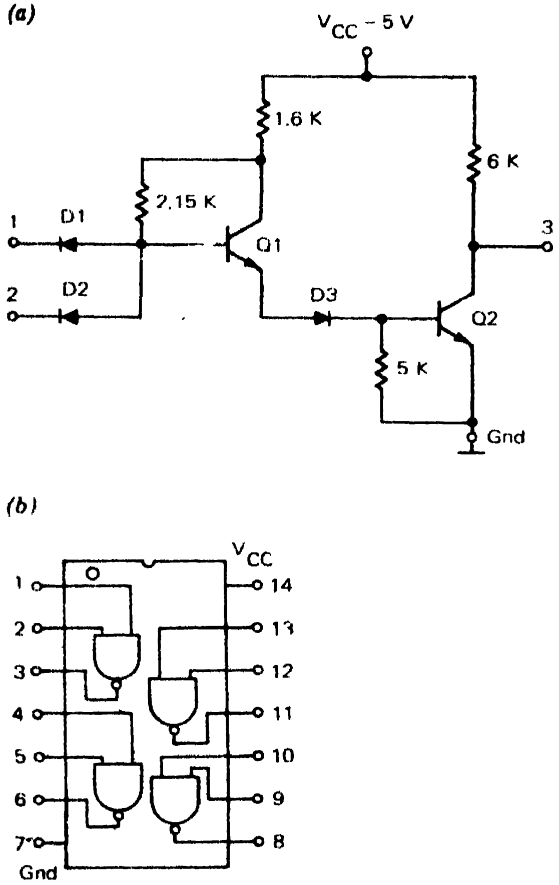

(a)
VCC - 5 V
1.6 K
6 K
2.15 K
1
D1
Q1
3
2
D2
D3
Q2
5 K
Gnd
(b)
VCC
1
2
3
4
5
6
7
Gnd
O
14
13
12
11
10
9
8

Fig. 25-9. (a) Schematic of a positive logic DTL NAND gate found in Motorola's MC846 package. (b) Showing the connections to the MC846 quad 2-input NAND gate. (Courtesy of Motorola Inc.)

It is left as a reader's exercise to show that the NAND gate combinations in Fig. 25-10 provide the AND and OR functions.

High Threshold Logic (HTL) NAND Gate

It was mentioned that diodes D3 and D4 in Figs. 25-8(a) and 25-9(a) were used to provide some noise immunity. Where a very noisy environment occurs, such as in industrial installations where motors and power lines are being switched on and off, even greater noise immunity is required.

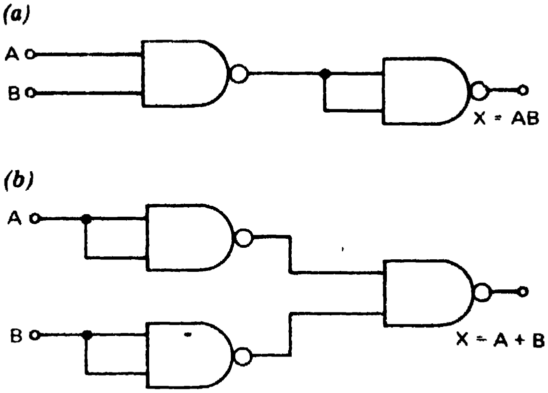

(a)
A
B
X = AB
(b)
A
B
X = A + B

Fig. 25-10. NAND gates used to form English logic: (a) AND gate. (b) OR gate.

This is achieved in the HTL gate in Fig. 25-11, where diode D3 in Fig. 25-9(a) has been replaced by a 7-V Zener diode. (Refer to Fig. 24-15(a) to recall how this is accomplished in an IC.)

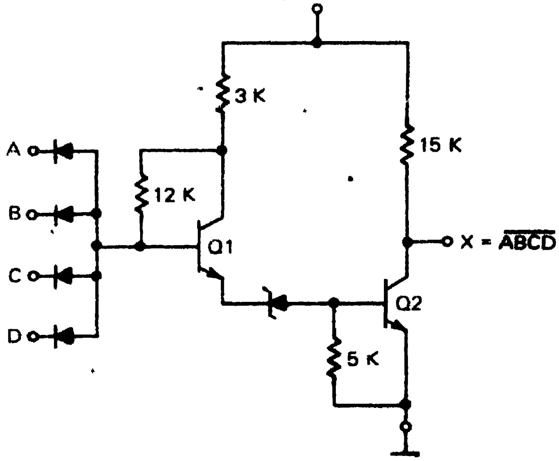

VCC = 15 V
3 K
15 K
A
B
C
D
12 K
Q1
X = $\overline{ABCD}$
Q2
5 K

Fig. 25-11. Example of a positive logic HTL 4-input NAND gate.

The use of the Zener diode means that the supply voltage must be raised and may be operated between 15 V and 18 V. The collector resistances are also increased. In this way, a noise immunity of approximately 6 V results. In fact, HTL is sometimes referred to as HNIL (high noise immunity logic), and the noise immunity may be given in terms of energy levels. The noise immunity energy of a logic device is a product of its voltage threshold, minimum propagation time (how long it takes a signal to move through the gate), induced current, and its output impedance. A typical HTL gate has a noise immunity energy of 5×10^{-9} joules, five times as high as its DTL counterpart. HTL is however restricted in the speed at which it can operate (this actually accounts for its high noise immunity level), so it is limited to industrial applications and not used in digital computer circuits.

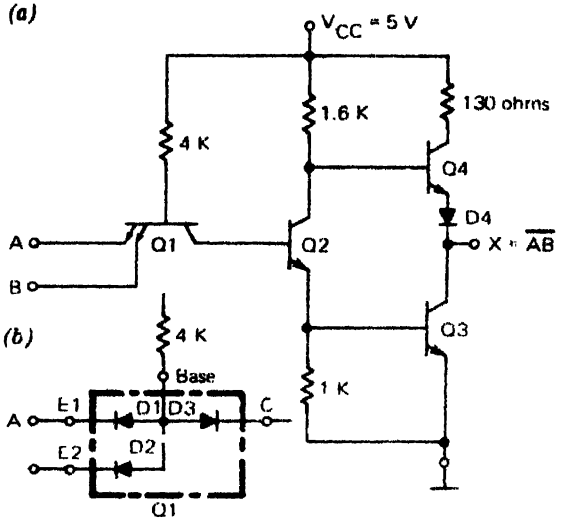


Fig. 25-12. (a) Schematic circuit of a 2-input positive logic TTL NAND gate found in Texas Instrument's SN7400 package. (Courtesy of Texas Instruments Incorporated.) (b) Simplified equivalent circuit of the input transistor.

Transistor Transistor Logic (TTL) NAND Gate

This family is the most popular of all the IC families, accounting for almost 50 percent of the digital sales in logic circuits. This popularity is due to a very low propagation delay/power dissipation product. A typical TTL (or T^2L) NAND gate would have a propagation delay of 10 nsec and a power dissipation of 10 mW. (A DTL gate may dissipate less power but in general has a higher propagation delay. Nevertheless, it is the next most popular in logic devices.)

The higher operating speed of a TTL gate arises principally from the use of an input transistor with multiple emitters, as shown in Fig. 25-12. This is one gate of a quad 2-input positive logic NAND gate, as found in Texas Instrument's SN7400 package.

The multiple-emitter input transistor is easily fabricated in an integrated circuit by making numerous (up to eight) N^+-type diffusions into the P-type base and making metal connections to each. The simplified equivalent circuit of Q1 shown in Fig. 25-12(b) suggests a very close similarity to the DTL NAND gate in Fig. 25-8. The emitter-base junctions provide the input diodes D1 and D2, while the collector-base junction functions as D3. D4 in Fig. 25-8(a) is provided by the base-emitter junction of Q2, and Q3 acts as the output transistor, like transistor Q. The only notable difference is in transistor Q4 and diode D4, which take the place of the 2.2-K ohm resistor in Fig. 25-8(a) but with a much lower effective resistance as will be explained.

Consider that both A and B are 1 (back biasing D1 and D2), so that the voltage at the base of Q1 is sufficient to cause D3 (the base-collector junction) to conduct and make Q2 and Q3 to go into saturation. The output voltage is approximately 0.2 V, and X = 0. Now if either A or B (or both) go low (0.2 V), the voltage at the base of Q1 is approximately

0.8 V, while the voltage at the collector of Q1 is approximately 1.2 V (the sum of the saturated base-emitter voltages of Q2 and Q3). This causes a momentary condition of reverse bias from collector-to-base of Q1, and, with the forward bias from base-to-emitter, a condition of normal transistor amplification occurs, with a resulting large collector current in Q1. This current is in such a direction as to cause base current to flow *out* of Q2 and remove its stored charge (and that of Q3 also) very quickly, making Q2 and Q3 turn off very rapidly. The output voltage goes to typically 3 V, and X assumes the 1 state. It is this *transistor* action of Q1 that results in a higher speed than is found in the ordinary back-to-back diodes in DTL.

The gain in speed achieved by Q1 would be lost if in some way the relatively slow rate of change in output voltage (due to the capacitive charging that takes place when a load is connected) is not somehow overcome. That is, if some capacitive load, C_L, (which is inevitably present in any load) is driven by the output in Fig. 25-8(a), there will be some delay when the output changes from 0.2 V and increases toward V_{CC}. This delay, determined by $R_L C_L$, is due to the exponential charging of C_L and may be excessively large, because of the relatively large value of R_L. If R_L were reduced, the delay would also be reduced, but this would increase the dissipation of the gate due to increased current through R_L when in the 0 state. R_L, in this case, is known as a *passive pull-up*, since it primarily determines the rate at which the output voltage builds up to its required value.

Q4 and D4 in Fig. 25-12(a) act as an *active pull-up* circuit, in which the effective resistance determining the exponential charging of C_L when Q3 is turned off is much reduced—to a little more than 130 ohms. That is, Q2 acts as a phase splitter, so that, when Q2 is turned off, Q3 is also turned off, but Q4 is driven into saturation. Thus, the charging time constant of C_L is determined by the sum of the 130-ohm resistor, the collector-to-emitter saturation resistance of Q4, and the forward resistance of D4.

The 130-ohm resistor is used to limit the current drawn from the power supply during switching. Diode D4 reduces the forward bias that would occur in Q4 when Q3 is in saturation, thereby ensuring that Q4 is cut off at this time and reducing the large dissipation that would otherwise occur in the 0 state. The Q4-D4-Q3 arrangement is sometimes called a *totem-pole* amplifier, because of the way in which one component "sits" upon the other. The typical TTL gate exhibits a noise immunity of 1 V.

Resistor Transistor Logic (RTL) NOR Gate

All of the previous logic families were more suited for producing the positive logic NAND gate. We now examine a NOR gate which is implemented using only resistors and transistors. RTL once enjoyed a good portion of the digital market, but its lower speed capability and higher power dissipation have limited its use now, compared with TTL and DTL.

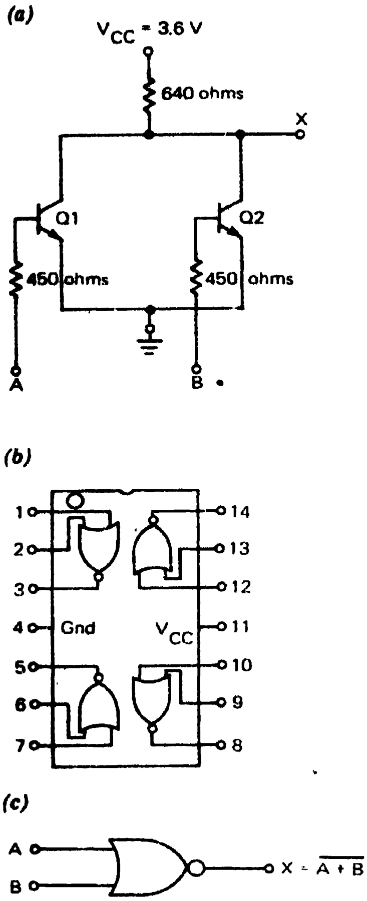
Fig. 25-13(a) shows the schematic of a positive logic RTL NOR gate which is one of four such gates contained in Motorola's MC724 quad 2-input NOR gate. The pin arrangement for the MC724 is shown in Fig. 25-13(b). Note that V_{CC} = 3.6 V and connections to V_{CC} and ground are different from those in the MC846 quad 2-input NAND gate of Fig. 25-9, which employs the same type of package.

The truth table indicates that an output results only if neither A NOR B is present. In Boolean form:

$$X = \overline{A + B} = \overline{A} \cdot \overline{B}. \qquad (25\text{-}13)$$

This may also be interpreted as saying X is high (1) if we have NOT-A AND NOT-B (both A and B = 0).

The operation of the circuit is quite straightforward. If both A and B are 0 (or floating),



(d)

Inputs		Output
A	B	X
0	0	1
0	1	0
1	0	0
1	1	0

Fig. 25-13. (a) Schematic circuit of a positive logic RTL NOR gate found in Motorola's MC724 package. (b) Showing the connections to the MC724 quad 2-input NOR gate. (Courtesy of Motorola, Inc.) (c) Standard symbol for the NOR gate. (d) Truth table for a NOR gate.

the transistors are cutoff, and the output voltage is somewhat less than 3.6 V, depending upon the connected load. Thus X = 1. If either A or B = 1, one transistor goes into saturation, and X = 0.

RTL uses very limited space, so it is economical to fabricate, but the noise immunity is the poorest of all the IC families.

It is left as a reader's exercise to show that substitution of NOR gates for NAND gates in Fig. 25-10 produces an OR function in (a) and an AND function in (b).

Direct-Coupled Transistor Logic (DCTL) NOR Gate

A DCTL circuit is the same as RTL circuitry with the base resistors (450 ohms) omitted. Thus, when one gate drives another, the collector of one gate is directly coupled to the base of the next gate. The circuit is usually operated at lower values of V_{CC}, for example 1.5 V, so that low breakdown voltage transistors may be used, and power dissipation is very small. However, the output voltage can only swing from 0.2 V to 0.8 V, so that noise immunity is very low. And the heavy saturation that each transistor experiences is detrimental to switching speed.

P-channel MOS Logic (MOS) NAND-NOR Gates

In MOSFET digital devices, only FETs are used; there are no resistors, diodes, or capacitors. Since the MOSFET requires a very small area for fabrication, this makes for a very economical implementation and is used widely in medium- and large-scale-integration applications. However, this high density of gates per chip makes it essential that the power dissipation per gate be very small. The high density is actually the greatest advantage of MOS logic, since its speed of operation is the slowest of all the IC families due to the inherent capacitances present at the inputs to the gates.

Fig. 25-14 shows how the NAND and NOR gates are constructed. Since the output of one gate (the drain) is coupled directly to the next gate with no intervening resistors, this is an example of DCTL. However, due to the insulated nature of the FET's gate at the input, excessive loading does not take place on the previous device's output, so that the disadvantage found in bipolar DCTL is not present here. It has a high fan-out capability.

Note that P-channel enhancement-type MOSFETs are used, so that the drain supply voltage V_{DD} must be negative. (Recall that 0 V at the gate of a P-channel MOSFET, with respect to the source, holds the transistor off, and a negative voltage—in excess of the threshold voltage—is necessary to turn it on.) In the gates shown, negative logic is normally used. That is, 0 V corresponds to a logical 0, and V_{DD} = -20 V to a logical 1. See Fig. 25-14(c).

It may be shown that Q1 in both circuits is always conducting, and its drain-to-source resistance acts as the load resistance (typically 2 K ohms) for the driver FETs, Q2 and Q3. Consider the NAND gate in Fig. 25-14(a). If either A or B (or both) is 0 (0 V), at least one of the series-connected, totem-pole-arranged transistors Q2 and Q3 will be off. A negligible current will flow through Q1, so that the output voltage is approximately -20 V. Thus X = 1, as required by the truth table in Fig. 25-8(c). Only if both A and B are present (1 = -20 V) will both Q2 and Q3 go into saturation and make X = 0.

In the NOR gate, the operation is quite the opposite, with a 1 at either A or B (or both) causing one of the parallel connected FETs to saturate and make X = 0. Only if both A and B are not present (0) will the output voltage go to -20 V and X assume the 1 state. This is as required by the truth table in Fig. 25-13(d).

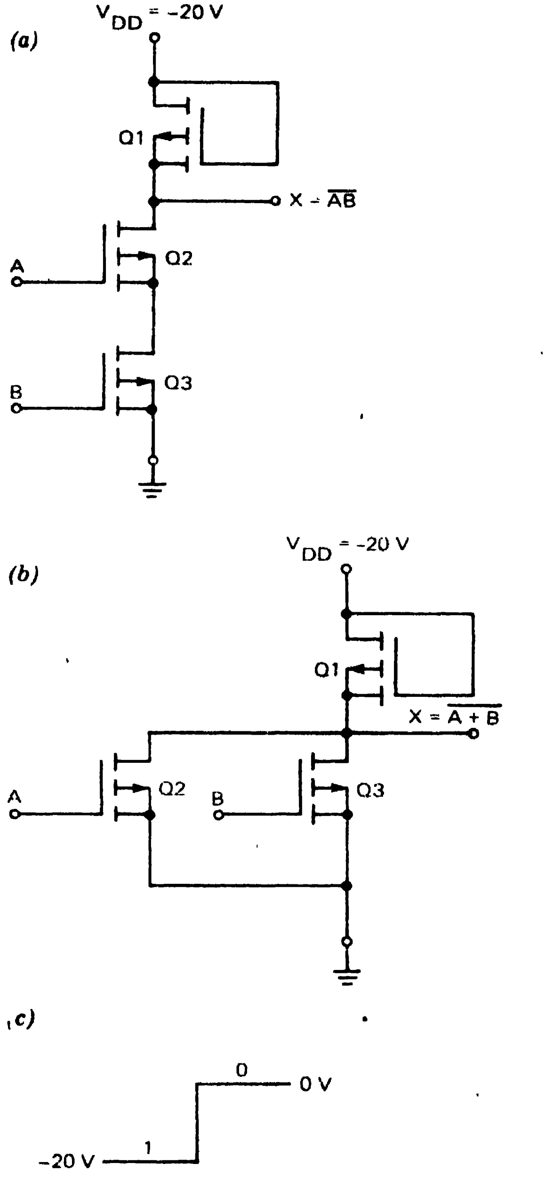


Fig. 25-14. (a) Schematic circuit of a negative logic MOS NAND gate. (b) Schematic circuit of a negative logic MOS NOR gate. (c) Showing the voltage levels for 0 and 1 for negative logic.

Complementary Symmetry MOS Logic (CMOS) NAND-NOR Gates

Although the power dissipation of the P-channel MOS devices is notably low, this may be reduced by a factor of 1000 in the complementary symmetry MOSFET devices. In addition, the operating speed is increased by at least a factor of four. The way in which this is achieved has been touched on in Section 24-5, where the basic inverter circuit was examined. Known as MCMOS by Motorola or COS/MOS by RCA, the complementary construction provides a method of reducing the quiescent dissipation in the on state (X = 0), by effectively shutting off transistor Q1 in the circuits of Fig. 25-14.

A CMOS NOR gate is shown in Fig. 25-15, along with a switching circuit to help explain its operation.

Note that this circuit uses positive logic, so that 0 V = 0 and +10 V = 1. Notice also the similarity with the MOS NOR gate in Fig. 25-14(b), as far as the parallel arrangement of Q2 and Q3 is concerned. However, the P-channel MOSFETs, Q1A and Q1B, do not act like linear resistors, but rather as switches, as shown in Fig. 25-15(b).

If A = B = 0 = 0 V, N-channel transistors Q2 and Q3 are off, but the gate-to-source voltage for P-channel transistors Q1A and Q1B is −10 V, causing them to saturate. Thus X = 1 = 10 V, as shown in Fig. 25-15(b). If either A or B (or both) = 1 = 10 V, either Q2 or Q3 (or both) will conduct, making X = 0. But, in addition, at least one of the transistors Q1A or Q1B will be cut off, because V_{GS} = 0 V. It is this feature that provides very little quiescent dissipation when the NOR gate is in the 0 state. Also, the turning on of one transistor, accompanied by the turning off of the other transistor promotes faster switching. The typical quiescent dissipation of the gate in Fig. 25-15 is 10 nanowatts, but its dissipation increases to 1 mW if the gate is operated at a switching speed of 1 MHz. Its noise immunity is typically ±4.5 V when operated at V_{DD} = 10 V, and its propagation delay is 75 nsec with a fan-out of 6. Delay increases directly with fan-out.

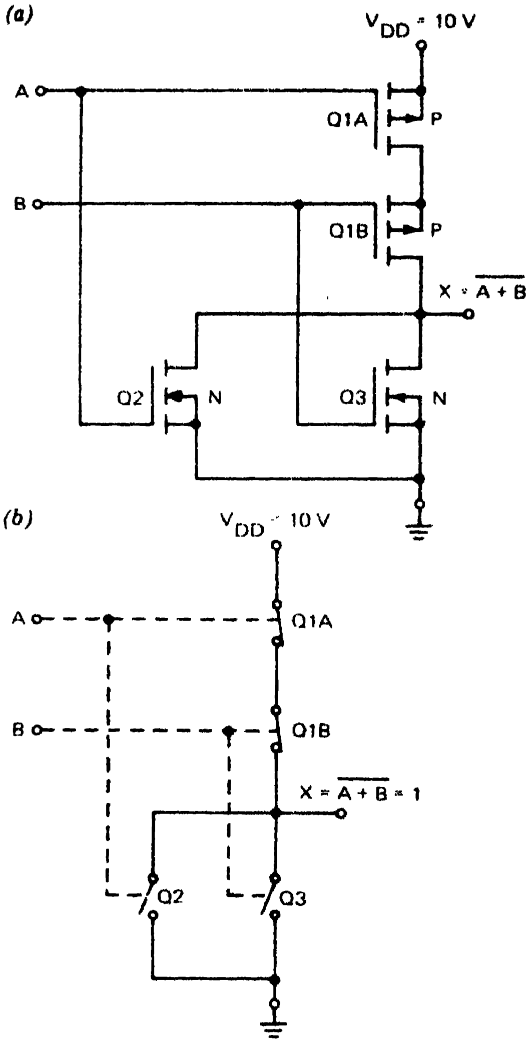


Fig. 25-15. (a) Schematic circuit of a positive logic CMOS NOR gate, part of a quad 2-input NOR gate found in RCA's CD4001 package. (b) Equivalent switching circuit for the NOR gate. Switches are shown for A = B = 0.

The NAND gate is just as easy to fabricate in CMOS logic as it is in MOS logic and has a similar arrangement, as shown by Fig. 25-16.

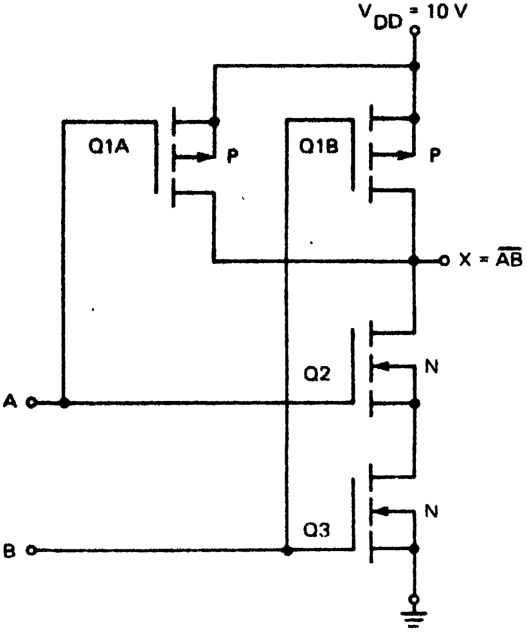


Fig. 25-16. Schematic of a positive logic CMOS NAND gate, part of a quad 2-input NAND gate found in RCA's CD4011 package.

Emitter-Coupled Logic (ECL) OR/NOR Gate

The remaining IC logic family differs from all the others in that the switching transistors are not allowed to go into saturation, although they do enter cutoff. In this way, the delay experienced in removing the charge stored in a saturated transistor is avoided, with the result that ECL gates exhibit the fastest switching speed of all the devices. However, this advantage is offset by the higher power dissipated in each gate.

The principle of ECL is found in the differential amplifier discussed in Section 23-9, as can be seen in the circuit of a basic ECL OR/NOR gate in Fig. 25-17. This is one of the dual 4-input gates found in Motorola's MC1204 package. As indicated by the symbol, both OR and NOR outputs are available.

Positive logic is used in this gate, with 0 = -1.55 V and 1 = -0.75 V. Note then that there is only a 0.8 V swing from one state to the other, which is much smaller than any of the other logic devices. In addition, observe that Q5, which forms one-half of the differential amplifier, has a bias voltage of -1.15 V applied to it. This value is exactly midway between the 0 and 1 values that are to be applied to the input transistors. This accurately controlled voltage may be supplied by a separate IC or contained within the same monolithic chip.

The differential amplifier is actually used as a comparator circuit. Recall from Section 23-9 that the operation of a differential amplifier is such as to maintain an approximately constant value of emitter current, I_E. This current is divided between the collector currents through R_{L1} and R_{L2}. In this circuit, operation is binary in nature, with the current flowing either through R_{L1} or R_{L2}. (In fact, ECL is sometimes called CML, current mode logic.)

Consider that A = B = C = D = 0 (-1.55 V). Transistors Q1 through Q4 are all off, while transistor Q5 is biased on, but it is *not* in saturation. This is because of the choice of R_{L2} and the value of I_E. The voltage drop across R_{L2} is approximately 0.8V, so that the base of Q7 is -0.8V with respect to ground. Q7 is in conduction, but being an emitter follower it is nowhere near saturation. Its emitter is 0.75 V lower than the base voltage, so that the voltage at X = -0.8 - 0.75 = -1.55 V, which is the logical 0, as required at the OR output.

Since no current flows through R_{L1}, the base and collector of Q6 are tied together, effectively making Q6 a forward-biased diode. Hence, the NOR output voltage at $\overline{X}$ = -0.75 V, which is the logical 1.

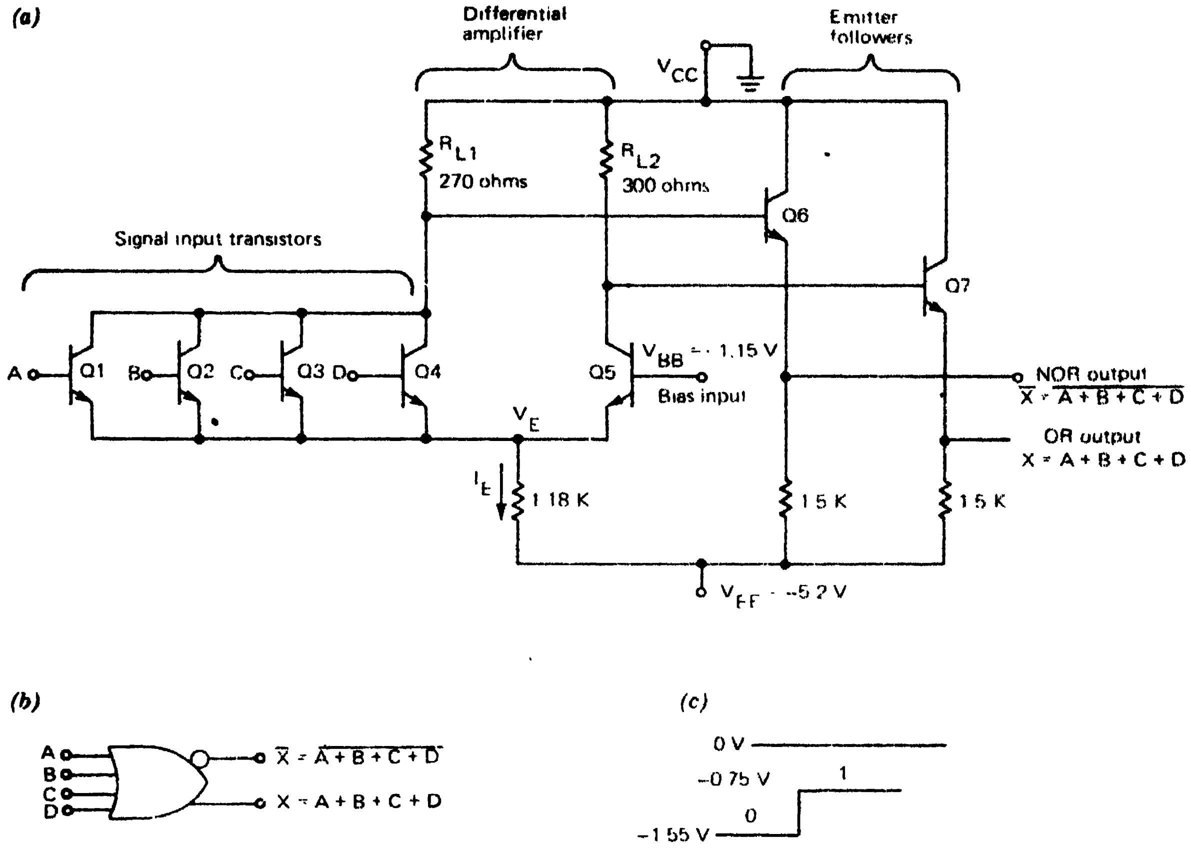

(a)
Differential amplifier
Emitter followers
V_CC
R_L1 270 ohms
R_L2 300 ohms
Q6
Q7
Signal input transistors
A
Q1
B
Q2
C
Q3
D
Q4
Q5
V_BB = -1.15 V
Bias input
V_E
NOR output X = A + B + C + D
OR output X = A + B + C + D
I_E
1.18 K
1.5 K
1.5 K
V_EE = -5.2 V
(b)
A
B
C
D
X = A + B + C + D
X = A + B + C + D
(c)
0 V
-0.75 V
1
0
-1.55 V

Fig. 25-17. (a) Schematic circuit of a positive logic ECL OR/NOR gate, part of a dual 4-input gate found in Motorola's MC1204 package. (b) Symbol for the OR/NOR gate. (c) Showing the logic voltage levels. (Courtesy of Motorola, Inc.)

If any input goes high (1 = -0.75 V), the bias input voltage is exceeded, one of the input transistors conducts, and current flow through the amplifier is switched from R_{L2} to R_{L1}. Transistors Q6 and Q7 assume the opposite state from before, with $\overline{X} = -1.55$ V = 0 and X = -0.75 V = 1.

The high input resistance offered by the differential amplifier and the low output resistance provided by the emitter followers make ECL have a very high fan-out capability.

25-5. COMPARISON OF LOGIC FAMILIES

Table 25-1 summarizes the important characteristics of the major IC digital logic families, assuming positive logic for all except MOS.

As can be seen from the table, various supply voltages and outputs are used, so that if different families are combined in the same digital system considerable *interfacing* may be necessary. The CMOS can be operated at a lower voltage, such as 3 to 5 V, to be compatible with DTL, TTL, and RTL, although reduced noise immunity then results.

TABLE 25-1 COMPARISON OF THE MAJOR IC DIGITAL LOGIC FAMILIES

CHARACTERISTIC	DTL	HTL	TTL	RTL	MOS	CMOS	ECL
Basic gate	NAND	NAND	NAND	NOR	NAND or NOR	NAND or NOR	OR/NOR
Fan-out (recommended)	8	10	10	5	20	50	25
Propagation delay, nsec	30	110	10	12	300	70	4
Power dissipation per gate, mW	11	44	15	25	10	1 μW quiescent 1 mW @ 1 MHz	25
Noise immunity	Good	Excellent	Very good	Average	Average	Very good	Excellent for power supply noise, low for signal noise
Power supply, volts	5	15	5	3.6	−20	10	−5.2 and −1.15
Logical unloaded 0/1, volts	0.2, 5	0.2, 15	0.2, 3	0.2, 3.6	0.2, −20	0.2, 10	−1.55, −0.75

25-6. IMPLEMENTING LOGIC CIRCUITS WITH NAND-NOR GATES

We saw in Section 25-3 how to use English (noninverting) gates to perform a required logic function. How is this done with NAND-NOR gates when the Boolean expressions are written in basic English logic? Consider the following example.

Example 25-3. Implement the following Boolean expression using (a) NAND gates, (b) NOR gates.

$$X = (\overline{A} + B)C + \overline{B}D + A\overline{C}\overline{D} + E. \qquad (25\text{-}14)$$

Solution. (a) (i) The first step is to draw the English logic diagram as in Fig. 25-18(a). (We will assume the inverted inputs are available.)

(ii) The logic diagram is redrawn with plain squares in place of the gates (since in some cases the resulting gates will be NORs instead of NANDs).

(iii) The inputs to these squares (gates) are modified according to the following rule: Count the number of gates from the output back toward the input. If the number is odd, negate the original input in the English logic diagram; if the number is even, leave the inputs unchanged. See Fig. 25-18(b).

(iv) If the last gate is an OR gate, as it is here, *all* the gates become NANDs. If the last gate is an AND gate, all gates become NORs.

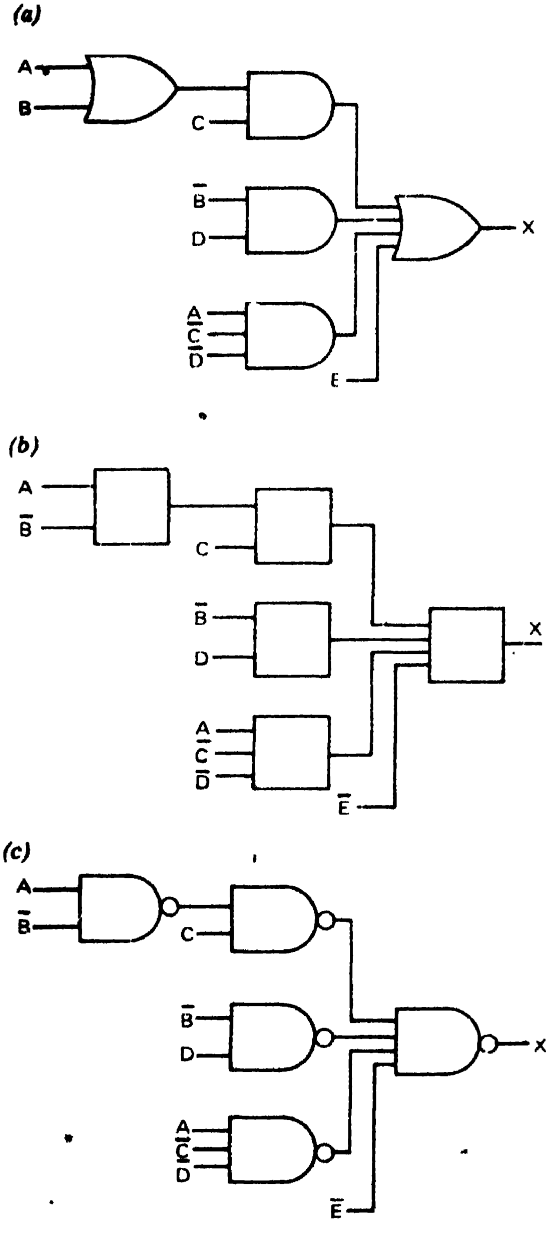


Fig. 25-18. (a) English logic implementation of Eq. 25-14. (b) Logic diagram showing the negated inputs for odd numbers of gates. (c) Completed NAND logic diagram.

The finished logic diagram is shown in Fig. 25-18(c) (The diagram in (b) may be omitted when it is known what the gates are changed to, NAND or NOR.) Although this may not always give the simplest configuration possible, this method is simple and saves time, which may be worth more than the saving of one or two gates. (A combination of NAND and NOR gates is sometimes used to provide an even simpler configuration. See Problem P25-17.)

(b) If we want the resulting gates to be NOR gates, it is plain that the last English logic gate must be an AND gate. To obtain this we must rearrange the given expression using De Morgan's Law:

$$X = (\bar{A} + B)C + \bar{B}D + A\bar{C}\bar{D} + E. \qquad (25\text{-}15)$$

Negate both sides of the equation:

$$\bar{X} = \overline{(\bar{A} + B)C + \bar{B}D + A\bar{C}\bar{D} + E} \qquad (25\text{-}16)$$

$$= (A\bar{B} + \bar{C})(B + \bar{D})(\bar{A} + C + D)\bar{E}. \qquad (25\text{-}17)$$

If we implement the right hand side of Eq. 25-17 in English logic, as in Fig. 25-19(a), the result is $\bar{X}$. We now redraw the logic diagram using NOR gates and invert the signals that pass through one or three gates, as shown in Fig. 25-19(b). A single-input NOR gate is now used as an inverter to produce X. Note that this is *not done* before counting the gates to determine which inputs should be inverted.

25-7. THE MEANING OF LOGIC IMPLEMENTATION

The reader may be left at this point wondering what the implementation of an equation such as Eq. 25-14 means. Although many complex logic situations arise in digital computer circuits, more readily understood ones occur in industrial applications. For example, in a nuclear power generating plant many safeguards and interlocks are used to ensure safe operation. Any one of the four terms in

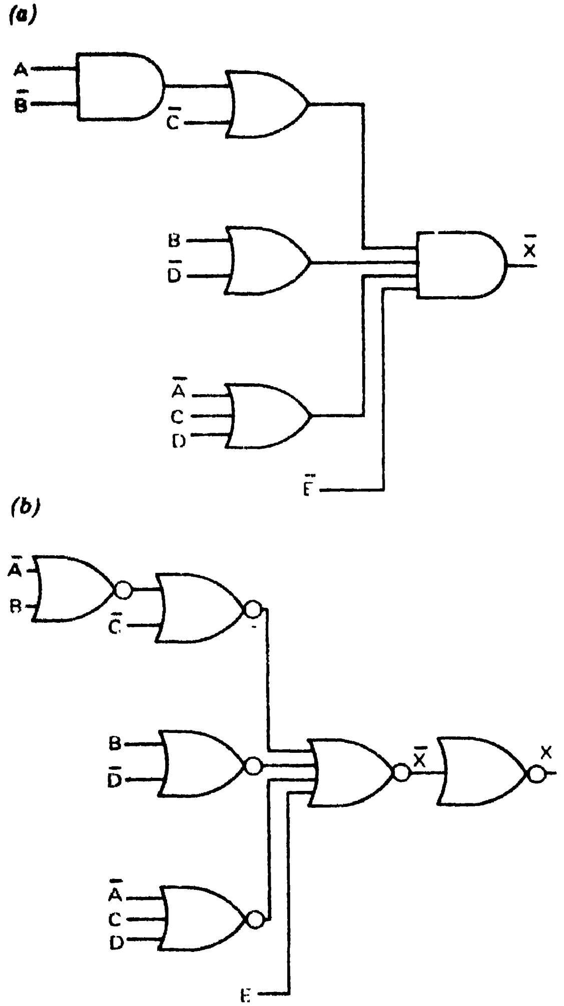


Fig. 25-19. (a) English logic implementation for $\bar{X}$ in Eq. 25-17. (b) Completed NOR logic diagram.

Eq. 25-14 may represent a condition to shut down the reactor. In some cases, a combination of inputs causes an unsafe condition. For example, the term $(\bar{A} + B)C$ may represent a condition in which (the main cooling pumps are not working ($\bar{A}$) OR the standby pumps are in use (B)) AND the line pressure has fallen below an acceptable value (C). Input voltages from the various transducers to indicate these conditions would make X go high and actuate a mechanism to lower the control rods and cease operation. In other cases, a single command such as an excessive temperature in the reactor (E) could trigger the same mechanism.

It should be mentioned that often in complex control applications such as in refineries, steel mills, power generating plants, etc., input voltages indicating conditions throughout the installation may be fed directly to a computer. These signals are applied to logic gates which have been *programmed* to take certain corrective action or make available printouts of the status of the variables at any requested time. This is called direct digital control (ddc) and requires that a memory function be provided for counting and decision making. We will look at some of the devices that provide this function in the next chapter.

Finally, we have discussed in this chapter primarily combinational digital logic *gates*. Also available in IC packages are combinations of these gates to perform definite functions, the most notable of which are found in the arithmetic section of a computer. These include full adders for binary addition (see Problem P25-18), decoders (BCD-to-decimal converters, etc.—see Problem P25-5 for BCD), and data selection. Time and space do not permit their consideration here, and the interested reader is referred to other books on digital principles and applications.

25-8. SUMMARY

(1) Combinational digital logic gates have an output at any given instant that depends only on the inputs that are present at that same instant. The order in which the inputs are applied is immaterial.

(2) The binary notation involves only two states, so that binary arithmetic only involves two numerals, 0 and 1, with the numeral 2 being the arithmetic base.

(3) In a positive logic level system, the more positive of the two voltage levels represents the 1 state, while in a negative logic system the more negative voltage represents the 1.

(4) English logic involves the basic AND, OR, and NOT gates, in which (a) an AND gate makes the output 1 when all the inputs are 1, (b) an OR gate makes the output 1 when one or more of the inputs are 1, (c) a NOT gate makes the output 1 when the single input is 0, and vice versa.

(5) AND and OR gates may be constructed from diodes and resistors called diode logic (DL). A disadvantage is the degradation of signal after passing through numerous gates.

(6) Boolean algebra is a method of writing logical expressions showing how an output depends upon the combination of inputs: a plus sign (+) between variables indicates the OR function, while a dot (•), or cross (×), or nothing, between the variables signifies the AND function. A line over the top of a letter represents inversion, or negation, or the NOT function.

(7) The EXCLUSIVE OR gate is a two-input circuit that provides an output when one input or the other is present, but not when both are present. Its Boolean symbol is a circled plus sign, ⊕.

(8) De Morgan's Law involves the negation of a whole Boolean expression: each input is negated, while AND gates become OR gates and vice versa.

(9) A NAND gate logically consists of an AND gate followed by an inverter. The output of a NAND gate is 1 if any of the inputs is 0; only if all inputs are 1 does the output go to the 0 state.

(10) A NOR gate logically consists of an OR gate followed by an inverter. The output of the NOR gate is 1 only if none of the inputs is present.

(11) There are seven major IC logic families whereby the NAND and NOR gates are constructed. These are summarized in Table 25-1. Except for MOS and CMOS logic, all use bipolar transistors. The MOS and CMOS gates have the greatest density per chip, due to the very small fabrication area needed for MOS field effect transistors. All the gates except ECL involve saturated transistors, so that ECL exhibits the fastest switching speed of all the IC logic families.

(12) A Boolean expression written in English logic is implemented using NAND or NOR gates by following the four steps in Example 25-3. Modification of the given expression using De Morgan's Law may be necessary to obtain the desired type of gate implementation.

QUESTIONS

Q25-1. In binary language, we often use 1 and 0 to represent the two states. Given one state as in the following table, provide the other for each condition.

ONE OF THE STATES	1	True		North	Hot	Off		
THE OTHER STATE			Low				Down	No

Q25-2. (a) What do the abbreviations LSB and MSB stand for?

(b) What is a *bit*?

(c) What is the base of binary arithmetic?

How does this explain why a number in binary is longer (has more digits) than its decimal counterpart?

Q25-3. Explain the difference between positive and negative logic, giving an example of each as applied to voltage levels.

Q25-4. (a) What basic gates constitute English logic?

(b) What does a truth table consist of?
(c) If a gate has 4 inputs, how many combinations of inputs are possible? What is a general rule to obtain the number of possible combinations for N inputs?

Q25-5. (a) In a DL AND gate, as in Fig. 25-3, is there any drain on the power supply when all three inputs are 0?

(b) If all three inputs go high, what does the output voltage depend upon?
(c) What is meant by *fan-out*?

Q25-6. Why is the unloaded output voltage of a DL AND gate 5 V, but only 4.4 V in a DL OR gate with the same voltage inputs?

Q25-7. (a) What does it mean to *negate* a signal?

(b) What logic "gate" is used to do this?
(c) If the input to the inverter in Fig. 25-5(a) is left "floating," what input does this represent and what is the output?

Q25-8. (a) How does an EXCLUSIVE OR gate differ from an INCLUSIVE OR gate?

(b) What is the Boolean equation for the output of an EXCLUSIVE OR gate, and what is its symbol?

Q25-9. What does De Morgan's Law apply to, and what does it state?

Q25-10. Does the output of a 2-input AND gate followed by one inverter differ from the output if the two inputs to the AND gate go through inverters first? Explain.

Q25-11. (a) Describe in your own words what takes place in the NAND gate circuit of Fig. 25-8(a) as the inputs change from A = 1, B = 0 to A = 1, B = 1.

(b) Describe the reverse process, as the inputs change from A = 1, B = 1, to A = 1, B = 0, indicating what effect a predominantly capacitive load will have on the time of response (the propagation delay).

Q25-12. Explain how HTL is essentially DTL modified to provide high noise immunity. Explain how the circuit provides this.

Q25-13. What two features in TTL provide higher operating speed than DTL, especially when working into the same capacitive load. Include the terms *passive* and *active-pull-up* in your discussion.

Q25-14. (a) What is the basic gate provided by the RTL IC family?

(b) If both inputs to the 2-input NOR gate in Fig. 25-13(a) are left floating, what input conditions does this represent, and what is the output?

Q25-15. (a) How does DCTL get its name?

(b) Why should the bipolar transistors in DCTL become more saturated and switch more slowly than in RTL?

Q25-16. (a) What single advantage does MOS logic have over all the bipolar transistor logic devices?

(b) What is its greatest disadvantage?

Q25-17. (a) How does CMOS differ from P-channel MOS logic in its quiescent dissipation?

(b) How is this achieved?
(c) Describe the operation of the CMOS NAND gate in Fig. 25-16, completing the truth table as you proceed.

Q25-18. (a) What is the distinguishing feature of ECL circuit operation compared with all the other logic families?

(b) What characteristic results from this feature?

PROBLEMS

P25-1. Convert the following binary numbers to decimal:

(a) 101, (b) 1111, (c) 10101101.

P25-2. Convert the following binary numbers to decimal:

(a) 11.1, (b) 1011.011, (c) 10010.1011.

(Hint: Digits to the *right* of the binary point correspond to increasingly negative exponents, 2^{-1}, 2^{-2}, 2^{-3}, etc.)

P25-3. Convert the following decimal numbers to binary:

(a) 16, (b) 23, (c) 53, (d) 101.

P25-4. Convert the following decimal numbers to binary:

(a) 5.5, (b) 53.8125, (c) 47.6.

(Hint: The decimal portion is converted separately by repeated multiplication by 2, the resulting 1 or 0 appearing in front of the decimal point becoming the binary portion, with the MSB appearing first.)

P25-5. For large numbers, true binary becomes very unwieldy. A more readily converted system is binary-coded decimal (BCD), in which 4-bit sets are used to represent units, tens, hundreds, etc. For example, the BCD form of 391 is: 0011 1001 0001. Convert the following to BCD:

(a) 23,
(b) 7,
(c) 101,
(d) 572,
(e) 1,984.

P25-6. Convert the following BCD into decimal:

(a) 0001 0101,
(b) 1000 1001 0010,
(c) 0001 1001 0111 0011.

P25-7. Classify the following as positive or negative logic:

(a) 0 = 0 V, 1 = 10 V;
(b) 0 = 0.2 V, 1 = -3.6 V;
(c) 0 = -1.5 V, 1 = -0.75 V;
(d) 0 = -0.8 V, 1 = -1.5 V;
(e) 0 = 5 V, 1 = 0 V;
(f) 0 = 3.6 V, 1 = 7.2 V.

P25-8. Given that positive logic is used, that is, 5 V = 1, 0 V = 0, convert the given circuit diagram into logic symbols and write the equation for the output, X. What type of gate is this?

P25-9. In the given circuit, A, B, and C represent the normally open or normally closed contacts that are actuated by relays A, B, and C.

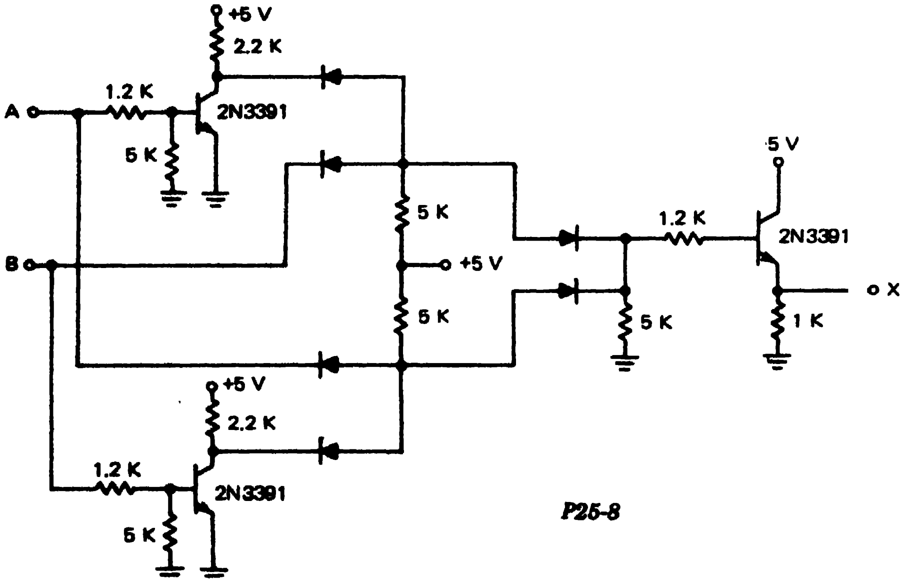
(a) Write the Boolean equation for the output.

(b) Simplify the equation given that $A\bar{A} = C\bar{C} = 0$ and draw the resulting circuit in relay form and logic gate form.

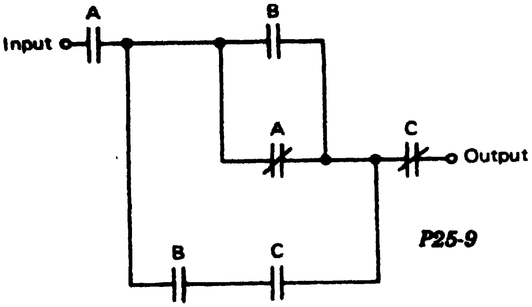
P25-10. Use a truth table to show that X $\bar{A}B + A\bar{B}$ and $X_2 = \overline{AB}\,(A + B)$ are equivalent. (Hint: For all the four combinations of A and B, obtain the states for X_1 and X_2 and compare. It may be useful to construct auxiliary columns for $\bar{A}B$, $A\bar{B}$; $\overline{AB}$, and $(A + B)$.)

P25-11. (a) Write the Boolean equation for X in the given logic diagram:

(b) Given that $C + \bar{C} = 1$, simplify the expression for X and determine what conditions (states) for A, B, C, and D will make X = 0.


+5 V
2.2 K
1.2 K
A
2N3391
5 K
5 V
5 K
1.2 K
2N3391
B
+5 V
X
5 K
5 K
1 K
+5 V
2.2 K
1.2 K
2N3391
5 K

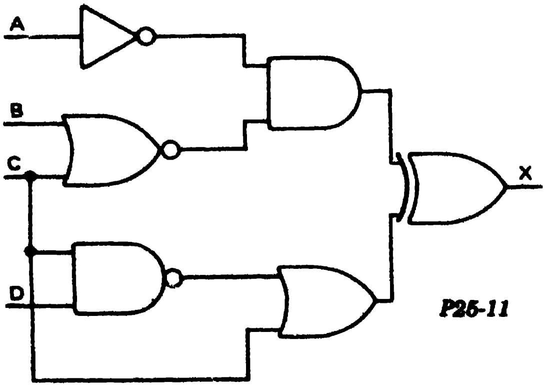
P25-8


A
B
Input
A
C
Output
B
C

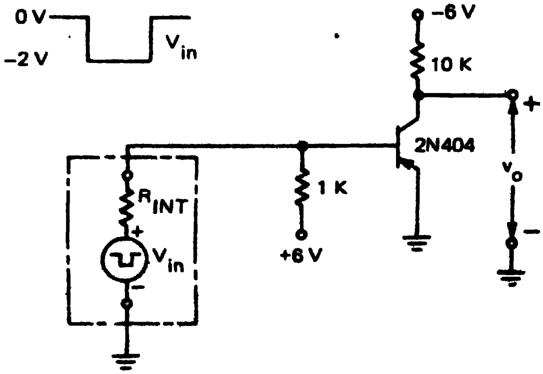
P25-9

P25-12. For the given circuit and V_{in} a negative pulse as indicated, determine v_o for the two conditions:

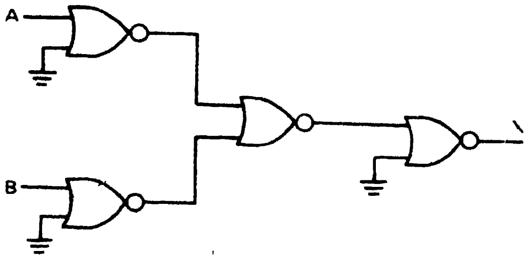
(a) $R_{INT} = 0$, (b) $R_{INT} = 1$ K ohm.



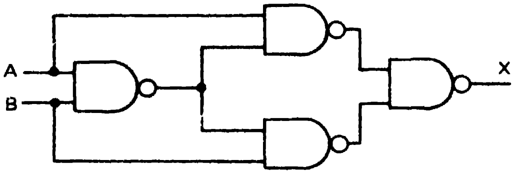
P25-11


0 V
−2 V
V_in
−6 V
10 K
+
2N404
v_o
−
R_INT
+
V_in
−
1 K
+6 V

P25-13. Write the Boolean equation for X in the given logic diagram and construct the truth table for it.



P25-14. Use the identity $\bar{A} + AB = \bar{A} + B$ to simplify the expression for the output X of the given logic diagram and state what type of function it is performing.



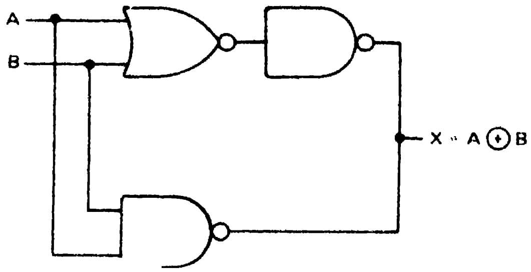
P25-15. Draw the logic diagrams to implement the following Boolean expression using (a) NAND gates, (b) NOR gates:

$X = \bar{A}(B + C) + D.$

P25-16. Draw the logic diagrams to implement the following Boolean expression using (a) NAND gates, (b) NOR gates.

$X = AB(C + \bar{D})(\bar{B}CE + \bar{A})F.$

P25-17. A method of implementing an EXCLUSIVE OR gate using a combination of NANDs and NORs is indicated. The connection between the two output NAND gates is called *wired-AND*, since it performs the AND function without using an AND gate.



Verify that the output $X = A \oplus B = A\bar{B} + B\bar{A}$ by using the identity $A\bar{A} = B\bar{B} = 0$ after invoking the algebraic distributive law. (Note how this circuit uses one gate less than all-NAND logic and three gates fewer than all-NOR logic.)

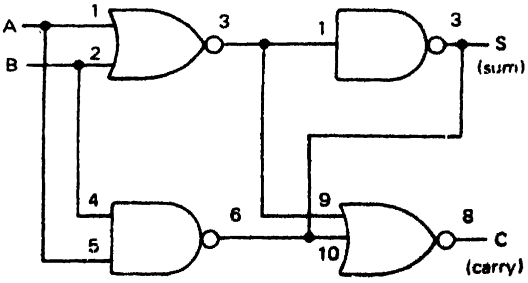
P25-18. A common combinational digital circuit that may be implemented using an MC724 NOR and an MC846 NAND is the half adder. It is used in binary arithmetic circuits to provide the "sum" and "carry" digits when two single digit binary numbers are added.

(a) For the given circuit, use De Morgan's Law and the distributive law as in algebra to show that the outputs at C and S are given by

$C = AB, \quad S = A \oplus B.$

(b) Construct the truth table for the circuit to verify that C = 1 and S = 0 only when the two inputs are both 1.

(01 plus 01 = 10.)



Note: Half adders are interconnected (with other gates) to form *full adders* that can handle three inputs, the two input digits and the carry from the previous stage. Complete 4-bit binary full adders are available in one package and represent an example of MSI since the number of gates is well in excess of 12; the number of components is in excess of 200.

26

NONLINEAR INTEGRATED CIRCUITS II: SEQUENTIAL DIGITAL CIRCUITS

26-0. INTRODUCTION

In this chapter, we consider an important branch of the multivibrator family, the *flip-flop*, since it forms the heart of all sequential digital circuits. The bistable multivibrator or flip-flop has two stable states and possesses the ability to stay in one of two possible states after an input has been applied. Since the state does not change following the removal of the input, the flip-flop is essentially a 1-bit memory or storage device. Many different forms of flip-flop exist, with the characteristics often depending upon the type of IC family that it belongs to. Most flip-flops are of the *clocked* type, in which the change of flip-flop state takes place at some definite rate. This keeps the signals moving through the digital system in step with each other (in synchronism), which is especially important in digital computers.

The other member of the multivibrator family, the monostable or one-shot, is found to have a useful function in some of the applications of flip-flops in registers and counters, as does the Schmitt trigger—a modified bistable multivibrator.

Emphasis in this chapter will be on functional block diagrams rather than on schematic circuitry, since the reader will probably be more concerned with using ICs as building blocks than with how they were made.

26-1. THE BASIC R-S FLIP-FLOP

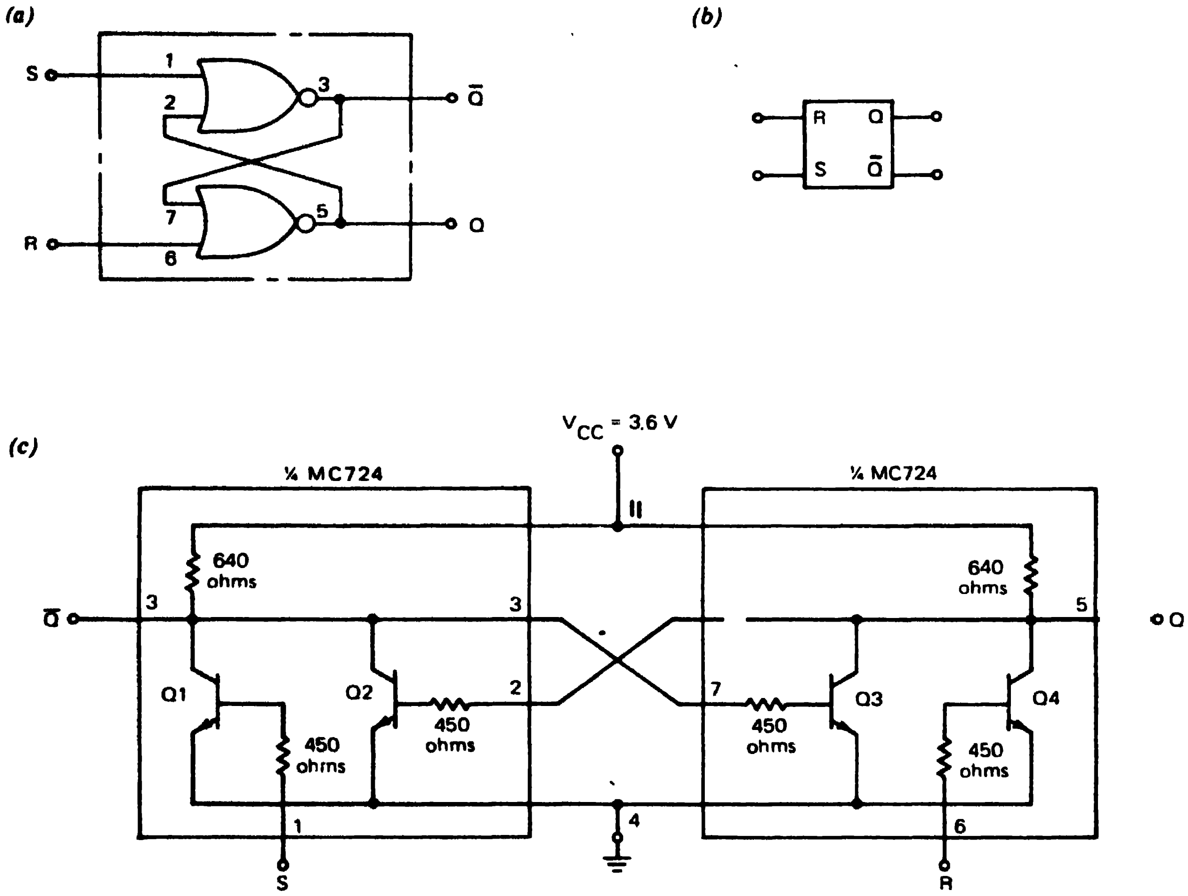
Consider the interconnection of two NOR gates as shown in Fig. 26-1. Pin numbers shown are for the MC724 quad 2-input NOR gate covered in Section 25-4.

Refer to the circuit in Fig. 26-1(c). Assume that, when power is first switched on and no inputs are applied, some circuit imbalance causes Q3 to go into conduction, and Q2 and Q1 are cut off. Using positive logic, $Q = 0$ and $\overline{Q} = 1$. The flip-flop is said to be in the *reset* position. Note that the positive voltage at the collector of Q2 is connected to the base of Q3, holding Q3 in conduction. This is one stable state.

Consider that inputs are applied so that $S = 1$, $R = 0$. This will cause Q1 to go into conduction, making $\overline{Q} = 0$. The removal of positive voltage from the base of Q3 (and the fact that $R = 0$) makes the collectors of Q3 and Q4 go high, so that $Q = 1$. The positive voltage at Q is fed back to the base of Q2, turning it on, so that the input to S may now be removed. The flip-flop is now in its other stable state, the *set* state, and will remain there if no further inputs are applied (or if the power is not interrupted). Of course, if $S = 1$, $R = 0$ is reapplied, no change will take place either.

Similarly, if now we make $S = 0$, $R = 1$, we will change states as shown by the truth table to *reset* the flip-flop. This may also be thought of as writing a 0 into the flip-flop since $Q = 0$; if $S = 1$, $R = 0$, we *write* a 1 into its memory, since $Q = 1$. Thus, a set input makes Q high, a reset input makes Q low.

If, simultaneously, the two inputs to the flip-flop are made 0, no change in the output states occurs. But, if both inputs are 1, both Q1 and Q4 go into conduction, so that $Q = \overline{Q} = 0$ for the duration of the input. As to which state the flip-flop will revert to will depend upon which input is removed first, or, if both are removed at the same time, what will happen will depend upon the circuit's characteristics. This is not desirable, so a 1-1 input is generally avoided.


(a)
S
R
1
2
3
5
6
7
Q̄
Q
(b)
R
Q
S
Q̄
(c)
V_CC = 3.6 V
¼ MC724
¼ MC724
11
640 ohms
640 ohms
3
3
2
7
5
Q̄
Q
Q1
Q2
Q3
Q4
450 ohms
450 ohms
450 ohms
450 ohms
1
4
6
S
R

(d)

R	S	Q	Q̄
0	0	N.C.	N.C
0	1	1	0
1	0	0	1
1	1	0	0

Fig. 26-1. (a) Two interconnected NOR gates form a basic R-S flip-flop. (b) Symbol of the basic R-S flip-flop. (c) Schematic circuit of the interconnected NOR gates. (d) Truth table of the basic flip-flop when constructed from RTL NOR gates.

A commercial example of the R-S flip-flop is the MC702 RTL package, which performs the equivalent of the circuit described. Note that one MC724 package can provide two R-S flip-flops.

26-2. CLOCK PULSES

It was mentioned that in many applications it is necessary for signals to move in synchronism throughout the digital system. The manner in which this is accomplished is by the use of a *master clock*. This essentially is an accurately controlled oscillator which produces a train of pulses, similar to that shown in Fig. 26-2. Whether the pulses are positive- or negative-going depends upon the system.

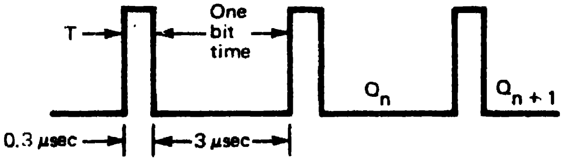


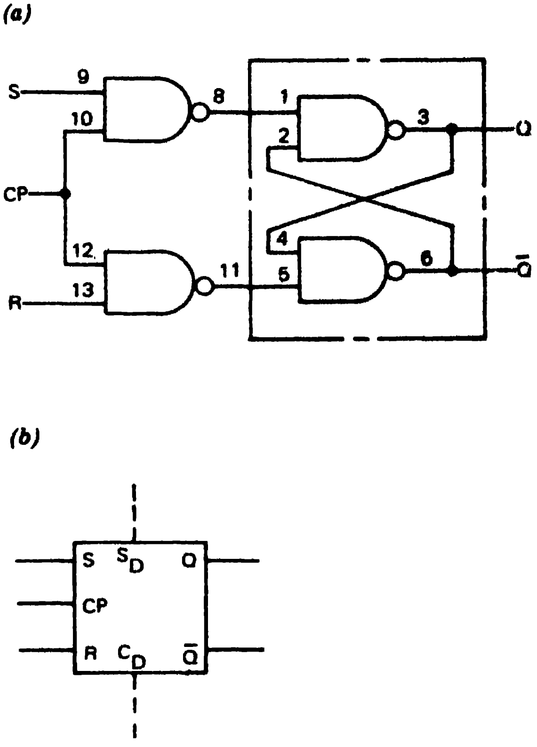
Fig. 26-2. A typical clock pulse (CP) waveform.

The duration of the pulse, T, is usually much smaller than the period of the clock rate, which in Fig. 26-2 is approximately 333 kHz. The time between clock pulses (CP) is known as a bit time and may vary from less than 1 μsec to over 15 μsec in large general purpose computers. This provides the time for certain operations to take place within the computer.

In clocked flip-flops, the desired change of state upon application of an input may only take place *during* the application of a clock pulse. (In some cases, the change of state only occurs on the rising or falling edge of the CP.) If, following one CP, a flip-flop's output is designated as Q_n, the state following the next CP is designated as Q_{n+1}. This notation will prove useful in constructing the truth tables of clocked flip-flops.

26-3. THE CLOCKED R-S FLIP-FLOP

A functional block diagram of a clocked R-S flip-flop is shown in Fig. 26-3 along with its truth table.



(c)

R	S	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	Q_n	$\bar{Q}_n$
0	1	1	0
1	0	0	1
1	1	Indeterminate	

Fig. 26-3. (a) Functional block diagram of a clocked R-S flip-flop (pin numbers shown for MC846). (b) Symbol for the clocked R-S flip-flop showing the direct set and clear inputs sometimes available (dotted). (c) Truth table for the clocked R-S flip-flop using DTL NAND gates.

It is seen that the circuit consists of a NAND gate R-S flip-flop whose inputs are gated (steered) by clock pulses fed into two additional NAND gates. Since inversion occurs in these input steering gates, Q now appears on the "same line" as S, and $\bar{Q}$ on the "same line" as R, as is clear from the symbol in Fig. 26-3(b). Also shown in the symbol are two additional inputs called the direct set (S_D) and direct clear (C_D), one or both of which are often made available. These have not been shown in the block diagram, but they permit the direct setting and clearing of the flip-flop without regard to the clock pulse input. Accordingly, they are called *asynchronous* inputs, since their effect on Q and $\bar{Q}$ does not depend upon the sequence in which they are applied, and they often override the synchronous inputs. (Note that the effect the S_D and C_D inputs have on the outputs depends upon the logic family. In DTL, typically $S_D = 1$ and $C_D = 0$ makes $Q = 0$ and $\bar{Q} = 1$, whereas in RTL, the same direct inputs make $Q = 1$ and $\bar{Q} = 0$.)

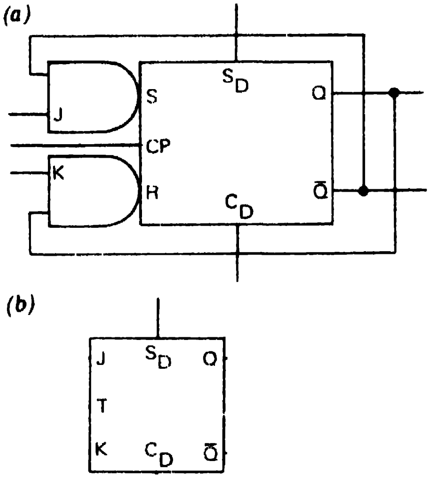
The truth table in Fig. 26-3 indicates the same type of operation as the basic R-S flip-flop in Fig. 26-1 except for the last row. (Note that using the notation Q_n, $\bar{Q}_n$, for $R = S = 0$ is equivalent to saying "no change" in the unclocked gate.)

The condition of the outputs for $R = S = 1$ has been labeled indeterminate for the following reason. The application of a CP when $R = S = 1$ makes the outputs from the first two NAND gates both 0. Consequently, both outputs from the second two NAND gates will be 1, indicating cutoff in the output transistors (positive logic). However, after the CP passes, both of the output transistors will try to go into conduction, with circuit imbalance determining the resultant state, since one transistor will go into saturation faster than the other. A condition of $R = S = 1$ must be avoided in R-S flip-flops, otherwise the outputs will be ambiguous.

26-4. THE CLOCKED J-K FLIP-FLOP—DTL, TTL

A method of avoiding the ambiguity of outputs from the R-S flip-flop for a 1-1 input condition is found in the J-K flip-flop. Actually the removal of the ambiguity is not the most important feature but rather the possibilities that are opened up in its applications. These will be considered later.

Fig. 26-4(a) shows the block diagram of a clocked J-K flip-flop. This is one of two found in the DTL MC852 dual J-K flip-flop package and is typical of those found in DTL and TTL clocked flip-flops.



(c)

J	K	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	Q_n	$\bar{Q}_n$
0	1	0	1
1	0	1	0
1	1	$\bar{Q}_n$	Q_n

Fig. 26-4. (a) Functional block diagram of a clocked J-K flip-flop. (b) Symbol for the clocked J-K flip-flop T for trigger or toggle is often used to designate the CP input. (c) Truth table for the clocked J-K flip-flop. Note the 1-1 inputs and that no change occurs if J=K=0.

Note that this J-K flip-flop can be identified as an R-S flip-flop with two AND gates at the inputs. The cross-connections from Q and $\bar{Q}$ to the K and J inputs gates, respectively, provide a means of steering a 1-1 input to the necessary S or R terminal to provide a change of state at the output. For example, consider that, prior to a CP, Q = 1 and $\bar{Q}$ = 0. The application of a 1-1 input at J and K will, following the next CP, make the R terminal go high (1), while S will be low (0). This, of course, makes $\bar{Q}$ = 1 and Q = 0, so that a change of state has occurred. See the truth table in Fig. 26-4(c). The state will change each time a CP appears if J = K = 1. It is this property of the flip-flop to *toggle* continuously from one state to the other at each CP that makes it very useful in counting applications. It explains the occasional use of T in the symbol in Fig. 26-4(b) instead of clock or CP; the J and K notations for the inputs were chosen arbitrarily but have since become standard.

26-5. THE CLOCKED J-K FLIP-FLOP—RTL

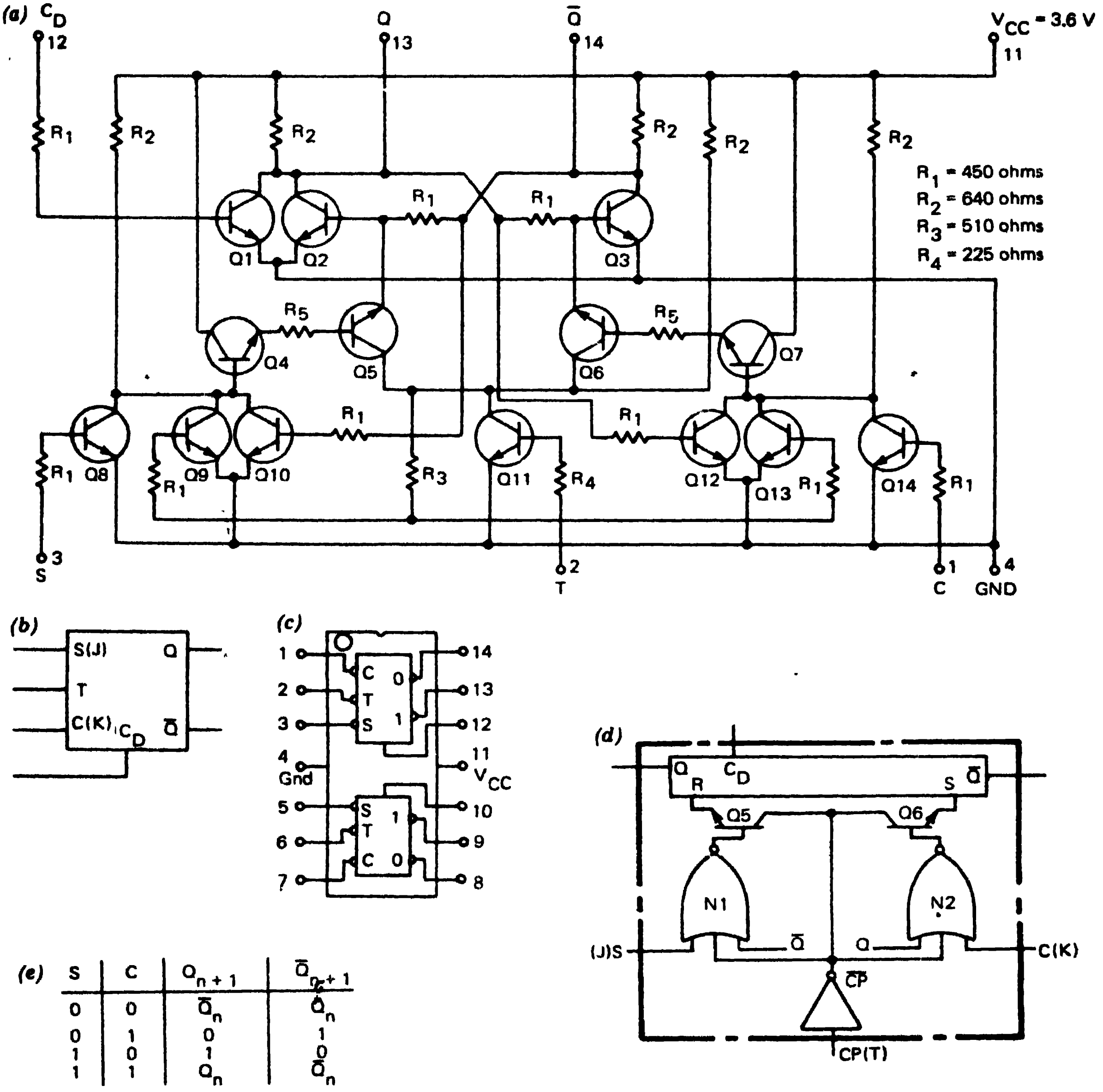
It may have been noticed that, instead of using additional AND gates at the inputs to the R-S terminals in Fig. 26-4(a), the cross-connected Q and $\bar{Q}$ inputs could have been applied to a *3-input* NAND in the R-S flip-flop itself. Refer back to Fig. 26-3(a).

A similar method to this is used in the MC790 clocked J-K flip-flop in the RTL family, a schematic circuit of which is shown in Fig. 26-5(a). This is one-half of a dual J-K package. However, as indicated by the symbol in Fig. 26-5(b), the inputs to this J-K flip-flop are usually labeled S and C (for *set* and *clear*). Note that only one direct asynchronous input is provided (direct clear, C_D), so that two of the flip-flops may be accommodated in one 14-pin DIP.

With reference to the circuit in Fig. 26-5(a), transistors Q2 and Q3 form a basic (NOR) R-S flip-flop. Trigger inputs to change the states of Q2 and Q3 are provided through the *steering* transistors, Q5 and Q6. These are selectively turned on by NOR gates, consisting of Q8, Q9, and Q10 on the one hand and Q12, Q13, and Q14 on the other. Transistors Q4 and Q7 act as drivers into Q5 and Q6, respectively. The whole circuit may then be redrawn in the functional block diagram of Fig. 26-5(d). Note that transistor Q11 provides an inverting function, so that if the CP arriving at T is as shown in Fig. 26-6(a), the waveform representing $\overline{CP}$ is as shown in Fig. 26-6(b). Thus, $\overline{CP}$ is held at 0 for the most part and goes to a 1 for a very short interval. It is this positive pulse directed through the appropriate steering transistor that is responsible for putting one transistor into conduction (output = 0), thus driving the other transistor into cutoff.

For example, say Q = 0 and $\bar{Q}$ = 1. Assume S = 1 and C = 0 are applied. Prior to the next clock pulse, $\overline{CP}$ = 0. Thus, all of the inputs to N2 are 0, making the output 1. This turns on Q6, and a path is *available* from the $\overline{CP}$ input terminal to the base of Q3 (the set input of the flip-flop). (Q5 is off.) Now, when $\overline{CP}$ goes to a 1, it can pass through Q6 and make Q3 go into saturation, which makes $\bar{Q}$ = 0. The cross-coupling to Q2 makes Q2 turn off, and Q = 1 as required by the set input. The positive voltage fed back to the base of Q3 takes the place of the $\overline{CP}$ input, maintaining a stable state. Note that when $\overline{CP}$ = 1 the output of N2 goes toward zero and ultimately turns Q6 off. However, the 1 input from $\overline{CP}$ has by this time already traveled through Q6 and performed its switching action.

It may be verified by a similar argument that the simultaneous application of S and C = 0 will, following the next CP, cause a change in state of the flip-flop to the opposite of what it was. It will continue to toggle as long as S


(a)
C_D 12
Q 13
$\bar{Q}$ 14
V_{CC} = 3.6 V
11
R_1
R_2
Q1
Q2
Q3
Q4
Q5
Q6
Q7
Q8
Q9
Q10
Q11
Q12
Q13
Q14
R_3
R_4
R_5
R_1 = 450 ohms
R_2 = 640 ohms
R_3 = 510 ohms
R_4 = 225 ohms
3 S
2 T
1 C
4 GND
(b)
S(J)
Q
T
C(K)
C_D
$\bar{Q}$
(c)
1
2
3
4 Gnd
5
6
7
14
13
12
11 V_{CC}
10
9
8
C
T
S
0
1
(d)
Q
R
C_D
S
$\bar{Q}$
Q5
Q6
N1
N2
(J)S
C(K)
$\overline{CP}$
CP(T)
(e)

S	C	Q_{n+1}	$\bar{Q}_{n+1}$
0	0	$\bar{Q}_n$	Q_n
0	1	0	1
1	0	1	0
1	1	Q_n	$\bar{Q}_n$

Fig.26-5. (a) Schematic circuit of one-half of the MC790 dual clocked J-K flip-flop. (Courtesy Motorola Inc.) (b) Symbol for one-half of the MC790. (c) Showing the pin arrangement for the dual MC790. (d) Functional block diagram of the clocked J-K flip-flop. (e) Truth table. Note that no change occurs if S = C = 1 (J = K = L) for RTL.

$= C = 0$ ($J = K = 0$). This is opposite to the clocked J-K flip-flops found in the DTL and TTL families. Similarly, if $S = C = 1$, there is no change of state following the CP, and the gate is said to be in the inhibit mode.

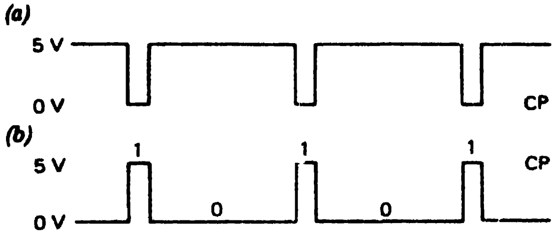


Fig. 26-6. (a) Waveform of negative going clock pulses, CP. (b) Waveform of positive going CP after passing through Q11 inverter.

26-6. OTHER TYPES OF FLIP-FLOPS

In some clocked J-K flip-flops, there may still be some ambiguity in the output when $J = K = 1$ is applied. This is because of a *race-around* condition in which the change in state of the fed back signals, Q and $\bar{Q}$, (to the input gates), that results when a CP appears, may cause an oscillatory condition for the duration of the CP. When the CP ends, the outputs may be quite arbitrary, depending upon the propagation delay through the flip-flop and the duration of the CP. (This particular problem is avoided in the circuit of Fig. 26-5(d), but it, too, may have problems, depending upon the condition of the CP, as mentioned in the next section.)

The J-K Master-Slave Flip-Flop

A circuit to avoid this race-around condition is called a J-K master-slave flip-flop and is shown in Fig. 26-7. It consists essentially of an input R-S flip-flop called the *master*, which is triggered by CP, and a tandem-connected output R-S flip-flop called the slave, triggered by $\overline{CP}$. The J-K inputs affect the master flip-flop when $CP = 1$, but, since $\overline{CP} = 0$, the slave is not excited, so its output state does not change during the clock pulse. When the CP has passed ($CP = 0$), the master is inhibited, but its output is now transferred to the slave, since $\overline{CP} = 1$. Erratic behavior does not occur during the clock pulse, since the cross-connected feedback is from the output of the slave to the input of the master. Since the outputs Q and $\bar{Q}$ do not change during the clock pulse, no oscillations can take place.

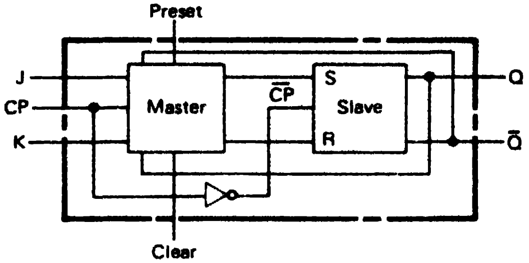


Fig. 26-7. Simplified block diagram of a J-K master-slave flip-flop.

The D-type Flip-Flop

If an R-S flip-flop is provided with an inverter on one of the input lines, only one data input to the flip-flop is necessary (D), as shown in Fig. 26-8(a).

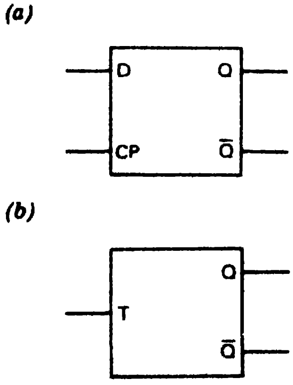


Fig. 26-8. (a) Symbol of a D-type flip-flop. (b) Symbol for a T-type flip-flop.

This is because the inverter ensures no ambiguous state (R = S = 1 is not possible). Direct set and clear inputs are sometimes provided. Also, the D input is often called the S input.

If a 1 is present at the D input (and Q = 0), Q does not become 1 until CP = 1. That is, the output after the CP equals the input at D before the CP. In this sense, the D-type flip-flop may be thought of as a *delay* (D) type of circuit, although this is true of any clocked flip-flop; change of state does not occur until (or after) the CP, even though the inputs may have arrived much earlier.

The D-type flip-flop is used as a temporary storage device and is known as a 1-bit bistable latch. The Monsanto MQL105 contains four such circuits and is known as a 4-bit bistable latch. If the clock input is held high (1), data at the D input are transferred to the Q output; if CP goes low (0), Q retains the information at the D input until once again the clock goes high. Then, if the input has changed, Q will change.

The T-type Flip-Flop

This flip-flop only has one active input, the clock or trigger input, T, as shown in Fig. 26-8 (b). The output changes state with each clock pulse applied at T, so it acts as a toggle switch. Although a quad type-T flip-flop is available (an MOS family MC1124L), this function is usually obtained from a J-K flip-flop. In this case, the J and K inputs are connected together to make J = K = 1 in DTL and TTL devices (see truth table in Fig. 26-4(c)). In RTL, it is necessary to make J = K = 0 (see Fig. 26-5(a)), which may be done by letting these inputs float, as in the MC790. It is this mode of operation that we will shortly explore further under the topic of counters.

26-7. CLOCK PULSE REQUIREMENTS

The propagation delay in the flip-flop—or the time taken for the flip-flop to change states—is typically less than 30 nanoseconds for the type of circuit in Fig. 26-5(a). If the fall time of the CP (or the rise time of $\overline{CP}$) is relatively large, say in excess of 100 to 200 nanosec, it is possible for the flip-flop to change state repeatedly during the same CP and give rise to false outputs. This is especially true if the trigger input to the flip-flop is derived (for some counter applications) from a metallic make-and-break switch, which often produces *contact bounce*. To avoid false triggering arising from these sources, the input trigger or clock pulses are *conditioned* to give fast rising or falling edges. This may be done by the use of a monostable multivibrator or a Schmitt trigger, both of which will now be described.

26-8. THE MONOSTABLE MULTIVIBRATOR (ONE-SHOT)

As indicated by its name, the monostable multivibrator has only one stable state and one quasi-stable (seemingly stable) state. It remains in its stable state until an input pulse triggers it into its other state. The duration for which it stays in this state is determined by the charging of an RC circuit. However, once the circuit returns to its stable state it remains there until another pulse arrives.

Fig. 26-9(a) and (b) show how the one-shot (since it provides one output pulse for every trigger input) may be constructed from two NOR gates, as found in the MC724 quad 2-input NOR gate. Elements R and C connected externally to the IC determine the duration of the output pulse and are selected for the particular application.

The operation of the one-shot may be followed either in the schematic circuit or in its equivalent NOR gate block diagram in Fig. 26-9(b).

When the switch is open (or no trigger input is applied), input 1 to Gate 1 is zero. Pin 7 on Gate 2 is high (3.6 V = 1), so pin 5 is low (0). Thus, the output of Gate 1 is high (3.6 V).

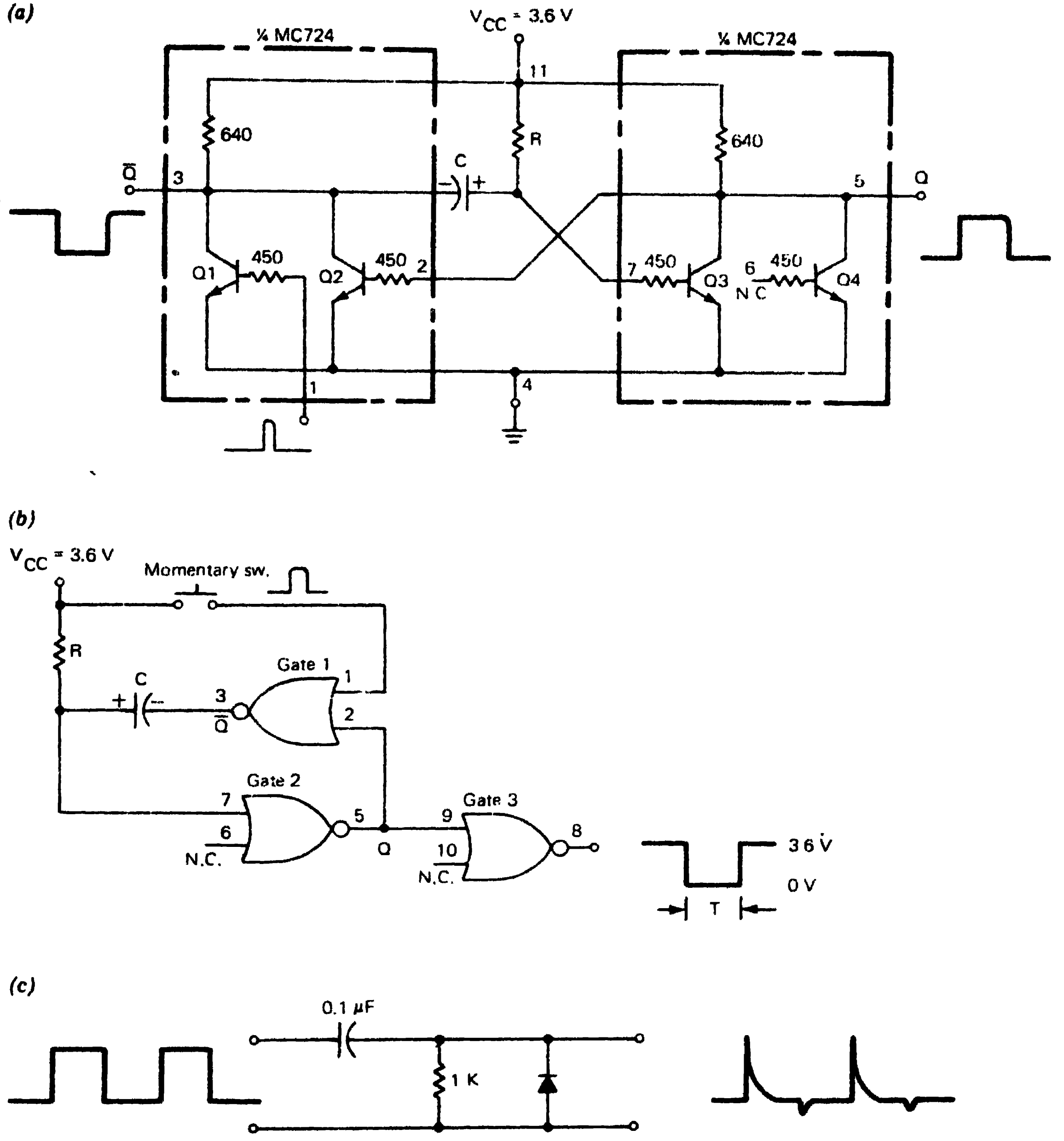

(a)
¼ MC724
V_{CC} = 3.6 V
¼ MC724
11
640
R
640
C
$\bar{Q}$
3
5
Q
450
450
2
7
450
6
450
Q1
Q2
Q3
N C
Q4
1
4
(b)
V_{CC} = 3.6 V
Momentary sw.
R
C
Gate 1
1
3
2
$\bar{Q}$
Gate 2
7
6
N.C.
5
Q
Gate 3
9
10
N.C.
8
3 6 V
0 V
T
(c)
0.1 μF
1 K

Fig. 26-9. (a) Schematic circuit of a monostable multivibrator using MC724 NOR gates. (b) Block diagram of the one-shot, showing how a momentary switch input can provide a fast falling edge at the output. (c) Differentiating circuit with a limiter to provide positive pulses from a square wave.

The capacitor is uncharged, since it has essentially 3.6 V on both plates. If the switch is momentarily closed (or if a positive pulse is applied to pin 1 of gate 1), pin 3 goes to 0, allowing C to charge through R. The removal of positive voltage from pin 7 makes the output of Gate 2 go high. This output voltage is fed back to Gate 1 and takes the place of the positive pulse. This switching action is regenerative, so that the rising edge of the output at pin 5 is very fast.

This condition is maintained until the voltage drop across R is reduced to a sufficiently small amount (due to C having become approximately 50% fully charged) so that the input to pin 7 of Gate 2 again goes high. Thus, the output at pin 5 drops to 0 but not as rapidly as the rising edge. (Why? See Section 18-11.)

The capacitor then discharges very rapidly (equalizing the voltage on its two plates), and the circuit is in readiness for another input. Note that if another trigger input had appeared while the circuit was in its quasi-stable state, no effect would be noticed, since the transistor in parallel with Q1 is already in a saturated state.

Fig. 26-9(c) shows a differentiating circuit, which may be used to produce positive pulses (from a square-wave input) suitable for triggering the one-shot. One may ask why should one go to the trouble of using a one-shot to produce a square-wave output if a square-wave input is already available? The answer is that the available square wave may not have a sufficiently fast rising edge to trigger some flip-flops. For example, the MC790 must be operated from a negative-going clock pulse with a fall time between 10 nsec and 100 nsec. The one-shot will provide this. (Master-slave J-K flip-flops do not require a very fast transition clock pulse. In fact, a 10-Hz sine wave may be used as a clock "pulse" to toggle some flip-flops of this type.) Gate 3 in Fig. 26-9(b) provides inversion, so that the edge with the rapid transistion is negative-going, as required by the MC790. Although the output at pin 3 provides inherent inversion, the waveform is modified by the action of the capacitor and would not normally be used unless it, too, were passed through an additional gate for shaping purposes.

An example in which the inverted output from the one-shot is used provides a second application of a one-shot. Consider that it is required to produce a pulse of a desired duration, T, delayed by some interval, D, following the appearance of an input pulse. This may be done using two one-shots and a differentiating circuit, as shown in Fig. 26-10(a).

The delay, D, will be determined by R_1C_1 in the first one-shot, and the duration of the output pulse, T, by R_2C_2 in the second one-shot. Waveforms are shown in Fig. 26-10(b). Note that one MC724 with two sets of external timing resistors and capacitors could provide the two one-shots.

The similarity of the one-shot in Fig. 26-9(a) to the flip-flop of Fig. 26-1(a) should be noted. When comparison is also made with the oscillator in Fig. 18-16, it is not hard to see how they all belong to the multivibrator family and how an astable multivibrator may be constructed using interconnected NOR gates with external resistors and capacitors.

26-9. THE SCHMITT TRIGGER

Another circuit to provide fast rising edges is the Schmitt trigger. A member of the multivibrator family, a Schmitt trigger (named after its inventor) is thought by some to be a modified one-shot. However, as will be explained, it more properly should be referred to as a bistable circuit. It is used to provide "squared" outputs from inputs that have a slow rise and fall time.

A simple way of constructing a Schmitt trigger from two interconnected NOR gates,

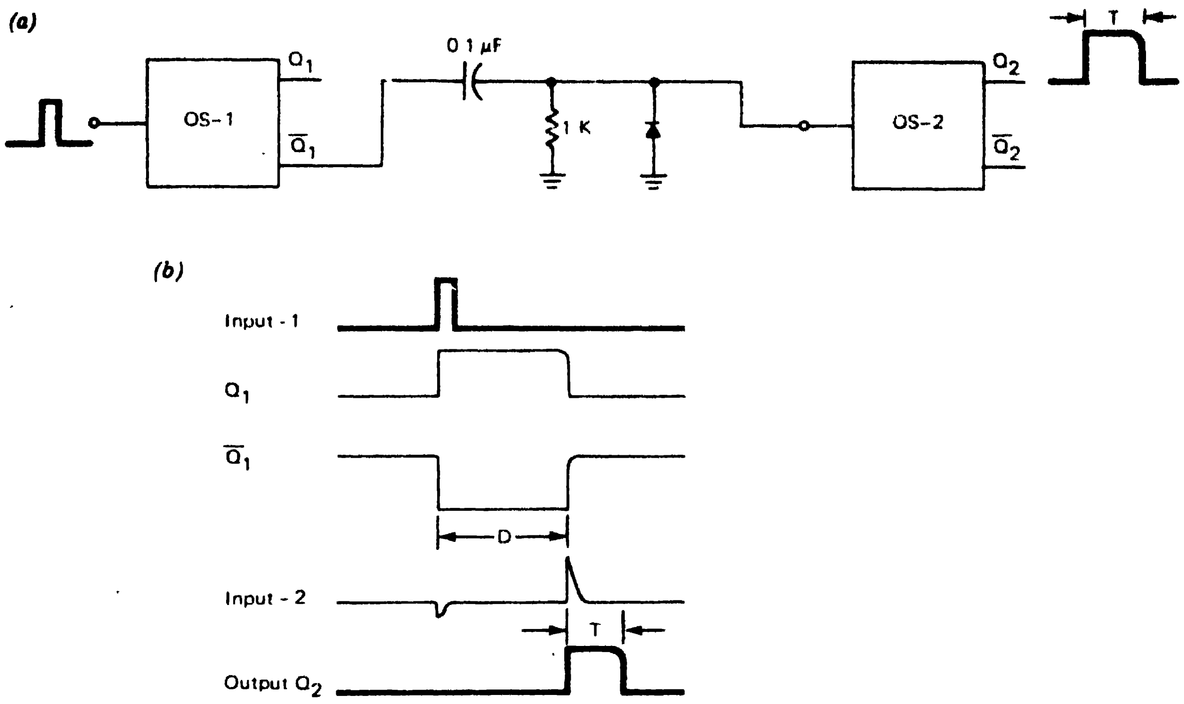


Fig. 26-10. (a) Circuit of two one-shots to produce a delayed pulse of desired width. (b) Waveforms.

as found in the MC724, is indicated in Fig. 26-11(a). As can be seen from Fig. 26-11(b), two additional resistors are needed, although somewhat inferior operation with R_e omitted is possible.

Consider for convenience the application of a sine wave, V_{in}. When $V_{in} = 0$ V, the output of Gate 1 is high, so that the output of Gate 2 is low. Put another way, in the quiescent state, Q1 is cutoff, so that Q2 is saturated. However, the output voltage is higher than the saturation voltage of Q2 (0.2 V) because of the voltage developed across R_e by the collector current through Q2. Approximately,

$$I_{C2} = \frac{3.6 - 0.2}{360 + 100} \approx 7.4 \text{ mA, and}$$

$$V_e \approx I_{C2}R_e \approx 0.74 \text{ V}.$$

Thus, the output is $0.2 + 0.74 \approx 1$ V.

Note that V_e reverse biases Q1, so that when V_{in} increases positively it must increase to approximately 1.4 V before conduction of Q1 will occur ($V_{BE} + V_e$). This voltage is called the input turn-on threshold voltage. A cumulative effect now takes place. As Q1 goes into conduction, the voltage coupled to the base of Q2 decreases, tending to cut Q2 off. The reduced current flow through Q2 and R_e reduces the reverse biasing effect on Q1 and aids in turning Q1 on. It is this regenerative action that results in a fast rising output pulse.

Q1 is now saturated, Q2 is cut off, and the output voltage is approximately 3.6 V. The output will remain in this state as long as the

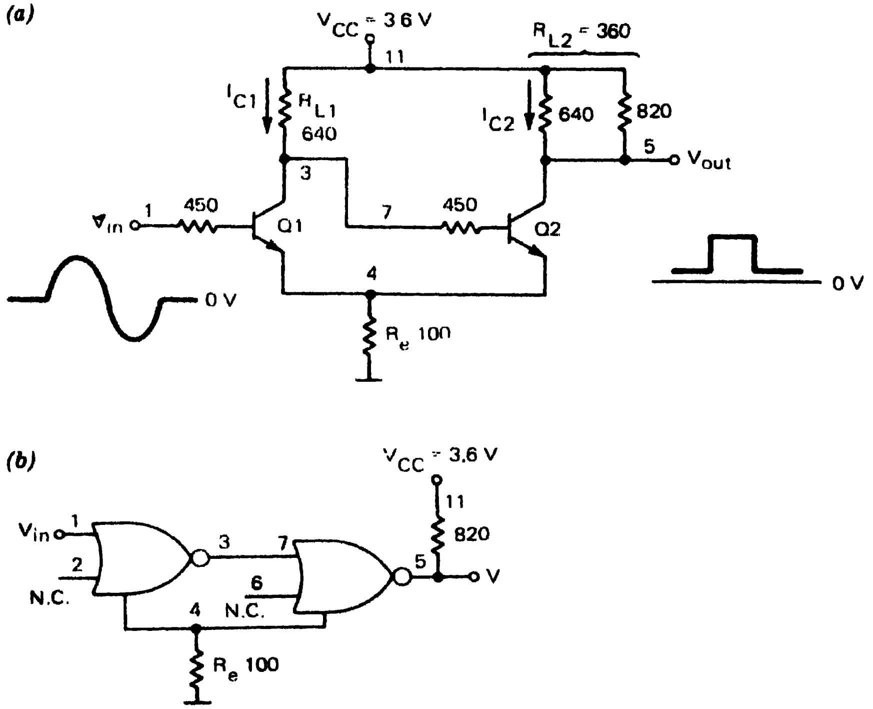


Fig. 26-11. (a) Circuit of a Schmitt trigger. (b) Showing how the Schmitt trigger may be constructed from two NOR gates.

input voltage exceeds the threshold voltage. Note that this is a stable state; it will not revert automatically to 0 at the end of some interval, as in a one-shot.

The collector current through Q1 (and R_e) is now approximately $I_{C1} = \frac{3.6 - 0.2}{640 + 100} = 4.6$ mA, and $V_e \approx I_{C1} R_e = 0.46$ V. Thus, as V_{in} decreases from its peak positive value, it must fall below 1 V approximately ($V_{BE} + V_e$) before Q1 will cut off. This is the turn-off threshold voltage. This condition makes Q2 conduct, driving the output down to its initial value, where it will stay even when the input goes negative. Not until V_{in} exceeds the input threshold voltage will the output go high again.

The different turn-on and turn-off threshold voltages at the input result in a ***hysteresis*** or ***backlash***, which may be shown graphically as in Fig. 26-12(a). (The amount of hysteresis may be controlled by changing the value of R_{L2} compared with R_{L1}.) The input turn-on threshold voltage may be increased by increasing R_e. Thus, the circuit may be used as a voltage comparator to provide an output whenever the input exceeds a desired quantity.

If R_e is omitted, two beneficial effects are lost. First, the regenerative action mentioned above to produce the fast rising and falling output is not available. Second, little or no hysteresis will result (there will be some hysteresis due to Q2 being in heavier saturation than Q1). Why is hysteresis desirable? As the input voltage, V_{in}, exceeds the input threshold voltage, base current into Q1 can reduce the value of V_{in} if it is obtained from a source of high internal resistance. If no hysteresis is present, the Schmitt trigger will immediately turn off again. But then V_{in} will again be high enough to start conduction, so oscillation (or "chatter") in the output will occur as V_{in} increases. In the circuit above, a 0.4 V margin is provided to avoid this problem.

The action of the Schmitt trigger to produce "squared" outputs from a slowly varying arbitrary waveform is shown in Fig. 26-12(b).

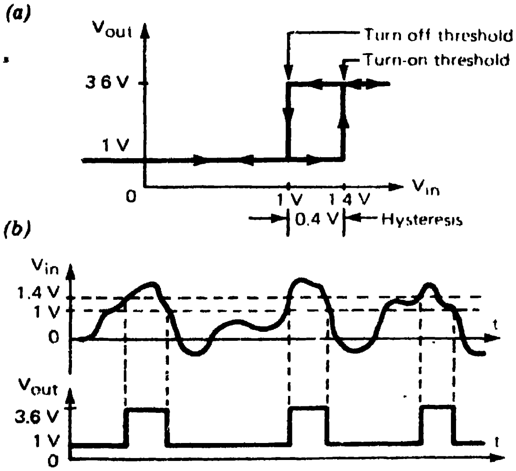


Fig. 26-12. (a) Graphical representation of hysteresis in a Schmitt trigger. (b) Showing the "squaring" action of a Schmitt trigger.

The Schmitt trigger may be used to advantage in combination with a one-shot. For example, if the input to be used with a one-shot has a slow edge or is just a level shift, a Schmitt trigger will produce the necessary square edge to effectively fire the one-shot.

Such a circuit is found in the SN74121 monostable multivibrator. This is a monolithic TTL IC that can produce output pulses varying in width from 40 nsec to 40 sec, by connection of external capacitance up to 1000 μF and external resistance up to 40 K ohms. The Schmitt input can accept a voltage that changes as slowly as 1 V/sec.

26-10. APPLICATIONS OF FLIP-FLOPS IN BINARY COUNTERS

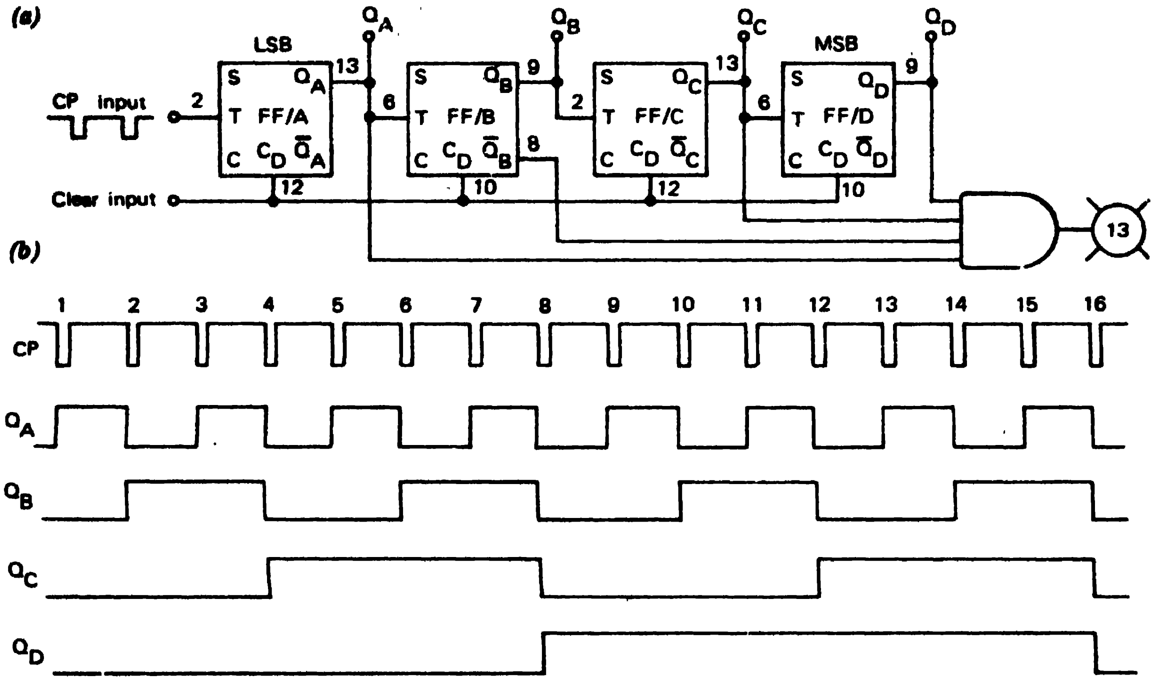
Many different kinds of counters may be constructed, and any attempt to cover them all here would do an injustice to the subject. Let us look at some of the more common ones available in IC form.

4-Bit Ripple-Through Counter

Consider four interconnected flip-flops, as shown in Fig. 26-13(a). The pulses to be counted (which may originate from a fixed-frequency clock pulse generator or be random in nature, as when counting objects on a moving conveyor belt) are applied to the first flip-flop. Its Q output acts as the trigger input to the next flip-flop and so on. The clock pulses shown are negative-going to satisfy the needs of the basic flip-flop found in the MC790 package. (Many flip-flops, especially the master-slave type, change state on the negative-going edge of positive clock pulses.)

In the RTL type of flip-flop (the MC790), the toggle mode requires that $J = K = 0$ ($S = C = 0$), which occurs when they are left floating. In other types, such as the DTL and TTL families, it is necessary to make $J = K = 1$ by connecting these to V_{CC}.

Assume that all flip-flops have been cleared (by a 1 on all the C_D inputs). After one CP, the output of A goes high. The second flip-flop is not affected by this positive-going edge. When the second CP arrives, A toggles, so that $Q_A = 0$. This falling edge triggers B, so that


(a)
LSB
MSB
Q_A
Q_B
Q_C
Q_D
CP input
Clear input
S
T
C
FF/A
FF/B
FF/C
FF/D
C_D
$\bar{Q}_A$
$\bar{Q}_B$
$\bar{Q}_C$
$\bar{Q}_D$
2
13
6
9
8
12
10
13
(b)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
CP
Q_A
Q_B
Q_C
Q_D

(c)

Input Pulses	Flip-Flop Outputs			
	Q_D	Q_C	Q_B	Q_A
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1
12	1	1	0	0
13	1	1	0	1
14	1	1	1	0
15	1	1	1	1
16	0	0	0	0

***Fig. 26-13.** (a) 4-bit ripple counter, showing how to decode a count of 13. (b) Waveforms of flip-flop outputs. Note the frequency division. (c) Table showing the state of the flip-flops at each count.*

$Q_B = 1$. It will remain in this state until two more CPs occur, since at that time A again produces a falling edge. The resulting change in Q_B from 1 to 0 now triggers C. It should be clear from the truth table and the waveforms in Fig. 26-13 how the CP inputs gradually "ripple through" the flip-flops.

After 15 clock pulses, all the flip-flop Q outputs are high. Thus, when the sixteenth CP arrives to reset A, all the others are also reset because of the falling outputs. The circuit then starts to count once more. The counter is said to have a modulus of 16 (2^4), since it goes through 16 unique combinations of outputs from the four flip-flops. (A ripple counter employing three flip-flops has a modulus of $2^3 = 8$.)

The circuit has two uses.

(1) As a counter, the outputs of the flip-flops after any desired count may be decoded by means of a logic gate to provide a visual indication of the count. For example, say that we would like to know when 13 clock pulses (or objects) have been counted. From the table in Fig. 26-13(c) at count 13, $A = 1, B = 0, C = 1, D = 1$. We could use a four-input AND gate with A, $\overline{B}$, C, and D as inputs. More specifically, we could use Q_A, $\overline{Q}_B$, Q_C and Q_D as inputs. Note that A is the LSB, and D the MSB.

(2) Examination of the waveforms in Fig. 26-13(b) reveals that *frequency division* has taken place. The frequency of the output from A is one-half that of the CP, while B's output frequency is one-quarter that of the CP, etc. By suitable modification and sufficient flip-flops, counters that provide "divide by 12," and "divide by 60" may be constructed. These are often used in digital clocks that are driven by the 60 Hz power line.*

A package that contains four master-slave flip-flops already interconnected to perform the above function is the SN5493. This is a monolithic MSI TTL high speed counter that can operate at an input count pulse frequency of 10 MHz.

A BCD Decade Counter

Since our natural counting system has a base of 10, it is convenient for many applications if a counter will reset itself to zero after a count of 10. One possible method of doing this is to decode a count of 10. That is, apply the output of an AND gate when $Q_B = 1$ and $Q_D = 1$ to the direct clear input, C_D. This, however, may not be reliable, due to the different delays in each flip-flop. Although a latch of the type mentioned in Section 26-6 could be used (at the output of the AND gate) to eliminate this problem, a solution that requires less gates is indicated in the binary coded decimal (BCD) counter in Fig. 26-14. This progresses through a binary count (and resets after a decimal count of 9), as does the ripple counter. It is called BCD or an "8-4-2-1 counter" to differentiate it from some counters that provide the same count but use forced feedback to skip some of the natural binary counts. The output weights of the flip-flops in these counters are not in accordance with the 8-4-2-1 code.

In the circuit shown, forced feedback is used, but it returns the counter to a reset state after nine natural binary counts. The truth table for this decade counter is thus the same as that in Fig. 26-13(c) up to and including count 9. On count 10, all Q outputs of the flip-flops are 0.

The circuit's operation may be verified by referring to the truth table. Briefly, A is seen to toggle continuously, so it is driven only by the input pulses. C changes state every time B changes from a 1 to a 0, so its trigger input is obtained from the output of B. B changes

---

*See Malvino, A., and Leach, *Digital Principles and Applications*, McGraw-Hill, New York 1969, Section 9-5.

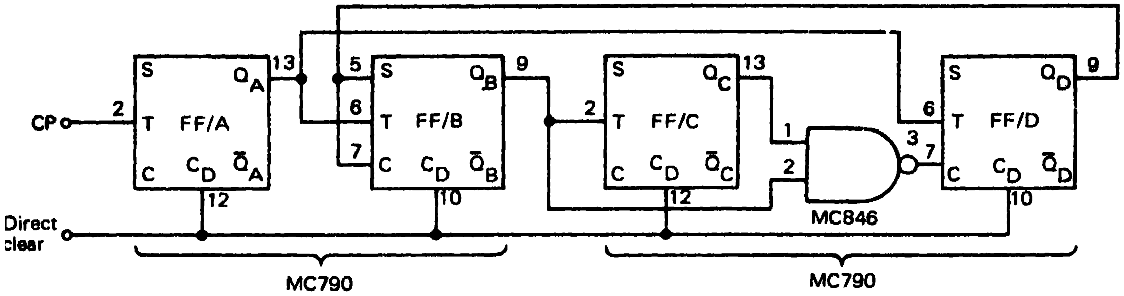


Fig. 26-14. A BCD decade counter constructed from 2 MC790 dual flip-flops and 1/4 MC846 quad 2-input NAND gate.

state every time A goes from 1 to 0 (so B can be triggered from A), *except* on the tenth pulse. B is inhibited from changing back to a 1 at this time by the 1 output from D being fed back to both of the main inputs of B. This provides the inhibit mode. Finally, note that D, although driven from the output of A, never changed state during counts 1 through 7. This is because of a sustained 1 being present at its clear input, due to either B or C (or both) being 0 when A dropped from 1 to 0. This clear signal is removed following count 7, so that Q_D goes high to indicate 1 (decimal 8) *when* A goes from 1 to 0. Observe that even though the clear input to D is once again activated after count 8, it will have no effect on D until A again drops from 1 to 0. This occurs on the tenth count, and all flip-flops go to 0.

An IC package that provides a BCD decade counter function is the SN5490. It, too, is an example of MSI and belongs to the TTL family. (It contains two NAND gates in addition to what is shown in Fig. 26-14, for auxiliary control applications.)

Cascaded BCD Decade Counters

When the previous BCD decade counter counts its tenth input pulse, Q_D goes from 1 to 0. This negative-going edge, when coupled to the LSB of another BCD decade counter, can be used to start a count of tens. When this counter is full, the last flip-flop will change from 1 to 0 and may be used to drive a third cascaded counter to indicate hundreds, and so on. Three cascaded BCD decade counters could then indicate a maximum of 999 for example, using only 12 flip-flops. See Fig. 26-15.

The decoding of each counter to show the decimal count is usually done by one IC package that provides an output suitable for driving an illuminated display. This display, to be covered in detail in the chapter on "optoelectronics," usually consists of a seven-segment device, which can assume the 10 decimal numerals 0 through 9 (and even some letters). As indicated by Fig. 26-15, the decoder is usually driven by a 4-bit latch which stores the numeral to be displayed.

26-11. SOME BASIC APPLICATIONS OF COUNTERS

In many applications, it is necessary to *gate* a counter so that it will count for a prescribed period. This may be done using the scheme shown in Fig. 26-16.

The rising edge of the enabling gate input may be passed through a differentiator circuit to reset the counter. At the end of period

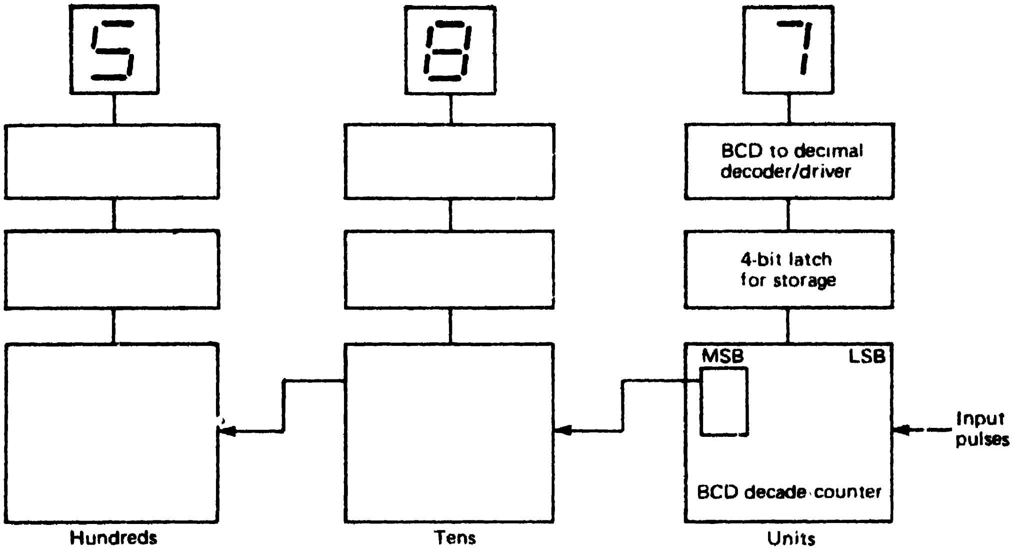

5
8
7
BCD to decimal decoder/driver
4-bit latch for storage
MSB
LSB
Input pulses
BCD decade counter
Hundreds
Tens
Units

Fig. 26-15. Showing how three BCD decade counters can be cascaded to provide a maximum readout of 999.

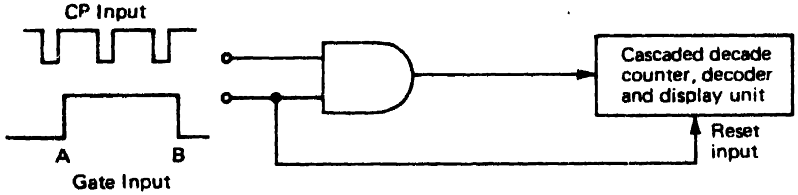

CP Input
A
B
Gate Input
Cascaded decade counter, decoder and display unit
Reset input

Fig. 26-16. Showing how to gate a counter for the period A-B.

A-B, the counter will retain the number of input pulses until again reset.

Principle of a Digital Frequency Counter

Figure 26-17 shows that the input pulses to the counter are obtained from the unknown signal frequency after passing through a Schmitt trigger for squaring purposes. If the interval A-B from the variable gate generator is exactly 1 second, the display indicated by the counter is directly the frequency. If the unknown frequency is changing rapidly, a smaller counting "window" (or shorter sampling period, say 0.1 sec) may be necessary to get an accurate frequency determination. In this case, the counter indication would have to be multiplied by 10, which could be done automatically by shifting a decimal point.

Principle of a Digital Voltmeter

In this case the variable to be measured, the voltage V, is applied to a voltage-controlled generator whose output pulse width W is proportional to V. See Fig. 26-18. This controls the period for which the counter

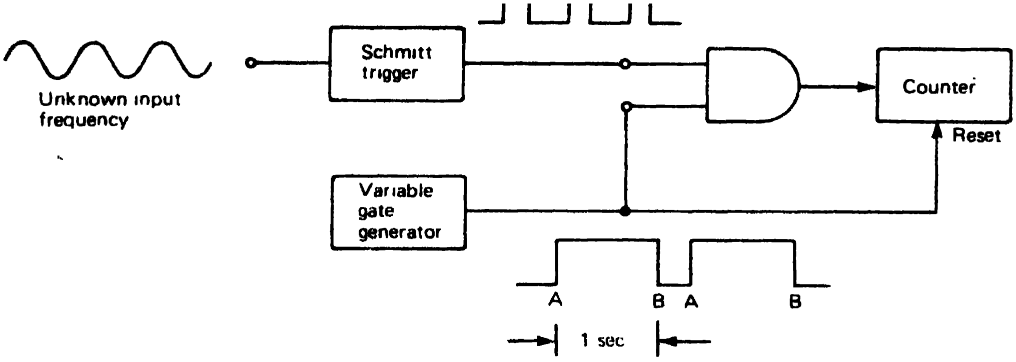

Unknown input frequency
Schmitt trigger
Counter
Reset
Variable gate generator
A
B
A
B
1 sec

Fig. 26-17. Circuit showing the principle of a digital frequency counter.

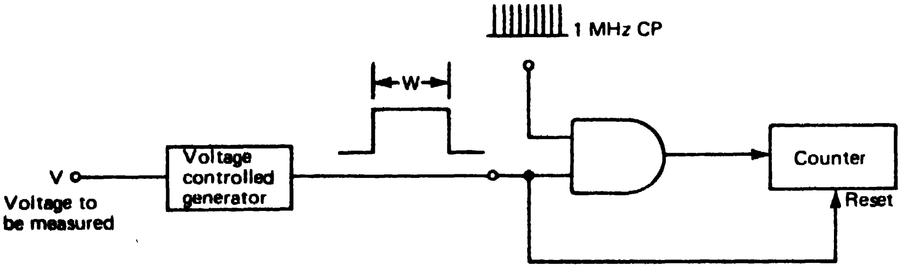

1 MHz CP
W
V
Voltage to be measured
Voltage controlled generator
Counter
Reset

Fig. 26-18. Circuit showing the principle of a digital voltmeter.

is permitted to count input pulses from an accurately known source, such as from a crystal-controlled 1 MHz clock pulse generator (oscillator). Thus, the counter indication is proportional to the voltage.

For example, say W is related to V by $W = V/10$ sec. Then, if $V = 2$ V, $W = 0.2$ sec, and the counter will indicate 200,000. The decimal point will, of course, be located after the first digit. If $V = 2.154$ V, then $W = 0.2154$ sec, and the counter will indicate 215,400. Evidently, the accuracy of the oscillator and pulse width generator, along with the number of cascaded decade counter units, will determine the overall accuracy and resolution of the voltmeter.

As a means of summarizing many of the logic concepts covered in the last two chapters, the following digital system is offered.

Left-Turn Traffic-Control System

Consider the problem of having to install a traffic light to permit traffic from a side street to make a left turn onto a heavily traveled main artery, as shown in Fig. 26-19(a). Inductive loops buried in the road at positions 1 and 2 provide an indication of cars (on the side street) approaching and leaving the intersection, respectively, by producing positive-going pulses.

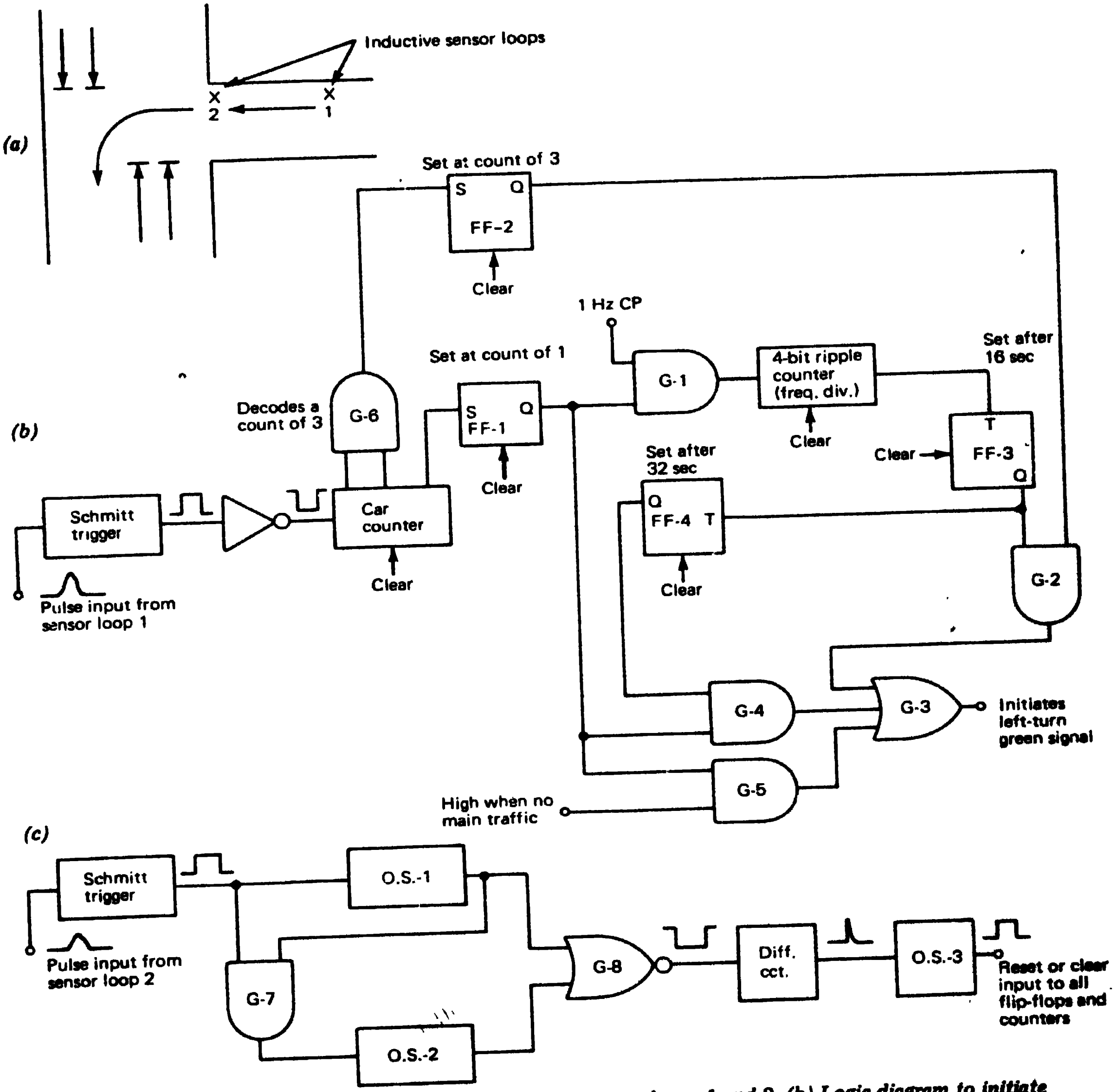

(a)
Inductive sensor loops
X 2
X 1
(b)
Set at count of 3
S Q
FF-2
Clear
1 Hz CP
Set at count of 1
G-1
4-bit ripple counter (freq. div.)
Clear
Set after 16 sec
Decodes a count of 3
G-6
S Q
FF-1
Clear
T
Clear
FF-3
Q
Set after 32 sec
Q
FF-4 T
Clear
Schmitt trigger
Car counter
Clear
G-2
Pulse input from sensor loop 1
G-4
G-3
Initiates left-turn green signal
G-5
High when no main traffic
(c)
Schmitt trigger
O.S.-1
G-8
Diff. cct.
O.S.-3
Pulse input from sensor loop 2
G-7
Reset or clear input to all flip-flops and counters
O.S.-2

Fig. 26-19. (a) Showing the left-turn intersection layout and sensor loops 1 and 2. (b) Logic diagram to initiate the turn-green mechanism. (c) Logic diagram to clear the flip-flops and remove the left-turn green signal.

Fig. 26-19(b) shows that these pulses, after squaring, are applied to a "car counter," which sets FF-1 high for one car and FF-2 high for three cars. The presence of one car waiting at the intersection enables G-1 and the 4-bit ripple counter has 1 cycle per sec (1 Hz) pulses applied to it from a slow-running multivibrator. The counter acts as a frequency divider, producing a negative-going pulse after 16 sec. This toggles FF-3 to a high state. If, during this 16-sec interval, at least three cars have accumulated, G-2 will provide an input to G-3 to initiate the turn-green mechanism.

If, however, only one or two cars are waiting, their "priority" is not high enough to warrant stopping the main traffic. The ripple counter will reset itself, and at the end of another 16 sec will produce another falling edge. This will cause FF-3 to toggle back to a low state and in so doing will toggle FF-4 into a high state. The presence of one car has been stored by FF-1, so G-4 is satisfied, producing an output from G-3. Thus, if only one or two cars are waiting, they will have to wait a total of 32 sec. However, if at any time during the second 16-sec interval another car is counted, G-2 will be satisfied, and a green light will result. If at least one car is waiting, and there is no traffic on the main artery (for example at night), G-5 will go high, and the green light will result without the 16-second delay.

A method of turning off the green light after the side street cars have exited is indicated in Fig. 26-19(c). The first pulse from sensor 2 will cause the output of O.S.-1 to go high for a predetermined interval. Consequently, G-8 will go low for the same interval (assuming no further pulses from sensor 2). As the output of G-8 returns to its high level, a positive spike from the differentiator circuit will fire O.S.-3. This positive pulse may be used to reset or clear all of the flip-flops and counters. The outputs of any activated AND gates will go to zero, and the side street green light will be turned off.

For one car, the duration of the output from O.S.-1 is sufficient to let one car exit. If other cars are present, subsequent pulses occurring while O.S.-1 is on are directed to O.S.-2. The output of O.S.-2 will "overlap" that from O.S.-1, and vice-versa, maintaining the output from G-8 low until no further car pulse inputs occur at either O.S.-1 or O.S.-2. This has the effect of delaying the turn-off pulse until all cars have exited, subject to the restriction built into the circuit that will be considered in the questions.

It is not claimed that the circuit described is the simplest one that will work, but it is a solution that uses all-solid-state devices (no mechanical clock timers, for example). Its main purpose is to illustrate the use of combinational and sequential logic gates in a digital system, employing most of the circuits we have discussed.

26-12. SUMMARY

(1) The basic R-S flip-flop is a bistable multivibrator that may be set high or low, according to the inputs. No change in state will occur for a 0-0 input, while a 1-1 input causes both outputs to go low if the device is constructed from cross-coupled RTL NOR gates.

(2) A clock pulse (CP) is used to move signals through a digital system in synchronism. It consists of fast rising and falling edges and is used to trigger flip-flops.

(3) A clocked R-S flip-flop differs from the basic R-S type in that change of state occurs only during the application of the CP. The output for a 1-1 input is indeterminate when fabricated from DTL NAND gates.

(4) A clocked J-K flip-flop avoids the ambiguity in outputs arising from a 1-1 input; the flip-flop will always change states (toggle) following each CP in DTL and TTL devices.

(5) A clocked J-K flip-flop constructed from RTL gates produces no change of state when S = C = 1 (J = K = 1); this is called the inhibit mode. When S = C = 0 (inputs floating), the flip-flop is placed in the toggle mode.

(6) A J-K master-slave flip-flop is used for high-speed operation and avoids the *race-around* problem. The master flip-flop is set during the application of the CP, and its output is transferred to the slave flip-flop after the CP passes.

(7) A D-type flip-flop has only one input (besides the CP input). It may be used to transmit input data to the output when the clock is high but stores the input data at its output when the CP goes low. It may also be known as a 1-bit bistable latch.

(8) A T-type flip-flop has only one input, the clock or trigger input. Usually constructed by making J = K = 1 (in DTL and TTL) or J = K = 0 (in RTL), its output toggles from one state to the other following each clock pulse.

(9) The monostable multivibrator (one-shot) produces one output pulse when triggered by a suitable input pulse. The duration of the output pulse may be selected by R-C components and is not affected by any more input pulses until it returns to its stable state. It may be used to "condition" momentary switch inputs to make them suitable for triggering flip-flops.

(10) A Schmitt trigger is a modified bistable multivibrator that functions as a voltage comparator with hysteresis. Its output changes state when the input exceeds a set turn-on threshold voltage; the output does not return to its initial state until the input falls below the turn-off threshold voltage. It may be used to "square" slowly varying inputs.

(11) A ripple-through counter consists of cascaded flip-flops (usually of the J-K master-slave type) with the output of one flip-flop feeding the trigger input of the next.

(12) The number of unique combinations of the flip-flops during counting is called the *modulus* (MOD) of the counter; these may be decoded to display the count visually.

(13) A ripple binary counter also performs a frequency-division function, with the output of the first flip-flop having a frequency one-half of the input pulse rate, the second flip-flop one-quarter, etc.

(14) A binary coded decimal (BCD) decade counter counts binarily up through 9 then resets itself. It is a MOD-10 counter that requires 4 flip-flops. BCD decade counters may be cascaded to indicate any size of number.

(15) Basic applications of counters involve the *gating* of a counter to count for a prescribed interval. Using this technique, a digital frequency counter, digital voltmeter, and traffic control systems may be constructed.

QUESTIONS

Q26-1. How does a basic unclocked R-S flip-flop using NOR gates differ from one constructed with NAND gates, as far as its truth table is concerned?

Q26-2. If an unclocked R-S flip-flop using RTL NOR gates has a 1-1 input condition, what are the Q and $\overline{Q}$ states? If now both S and R are made 0 (at the same instant), what will happen? (Consider what is meant or implied by the phrase "at the same instant.")

Q26-3. Describe why a 1-1 input condition to a clocked R-S flip-flop using NAND gates causes an indeterminate output.

Q26-4. (a) What is a clock pulse used for?
(b) What is meant by a "bit time"?
(c) What does the notation Q_n, Q_{n+1} indicate when used with clocked flip-flops?

Q26-5. (a) What is an asynchronous input?
(b) What does it mean to directly set a flip-flop?
(c) Can a clocked J-K flip-flop be directly cleared without waiting for a clock pulse?

Q26-6. (a) How does a clocked J-K flip-flop differ from a clocked R-S flip-flop?
(b) How does a clocked J-K flip-flop in the RTL family differ from DTL or TTL J-K flip-flops, as far as the truth table is concerned?
(c) What is meant by the toggle mode? The inhibit mode?

Q26-7. Describe briefly how a J-K master-slave flip-flop avoids the *race-around* problem sometimes found in other types of J-K flip-

Q26-8. (a) What is a 1-bit bistable latch?
(b) What is it also known as?
(c) What is a T-type flip-flop?
(d) How may it be obtained from a J-K flip-flop?

Q26-9. (a) What is the duration of the output pulse from a one-shot dependent upon?
(b) How does this differ from the way in which an astable multivibrator stays in one of its unstable states?
(c) Why does another input pulse to a one-shot not affect the output if it has just previously been triggered?

Q26-10. (a) What is a differentiator circuit?
(b) Why is it so called?
(c) Explain in your own words how two one-shots can produce a delayed pulse having a desired width.

Q26-11. (a) What is the function of a Schmitt trigger?
(b) Why may a Schmitt trigger be thought of as a voltage comparator?
(c) What two factors determine the input turn-on threshold voltage?
(d) What factors determine the turn-off threshold voltage?
(e) What is meant by the hysteresis of a Schmitt trigger?
(f) Why may hysteresis be desirable?

Q26-12. (a) Why is the name *ripple* counter appropriate?
(b) What is meant by frequency division?
(c) What is the modulus of a five-flip-flop ripple counter?
(d) What frequency division would occur at the output of the third flip-flop?
(e) In a four-bit ripple counter, what Q inputs would be necessary to an AND gate to decode a count of 9?
(f) If the inputs to the counter originate from a mechanical toggle switch, under what conditions would you suggest using a one-shot for pulse conditioning?

Q26-13. (a) What is a BCD decade counter?
(b) Why is it also referred to as an 8-4-2-1 decade counter?
(c) What feature is used in a four-flip-flop decade counter to make it stop at 10 instead of 16?

Q26-14. (a) Explain how BCD decade counters may be cascaded to give any size of count.
(b) How is the output usually decoded to give a visual display of the count?

Q26-15. (a) What does it mean to *gate* a counter?
(b) Explain the principle of a digital

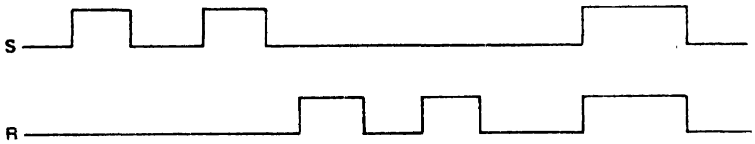
frequency counter.

(c) Explain the principle of a DVM (digital voltmeter).

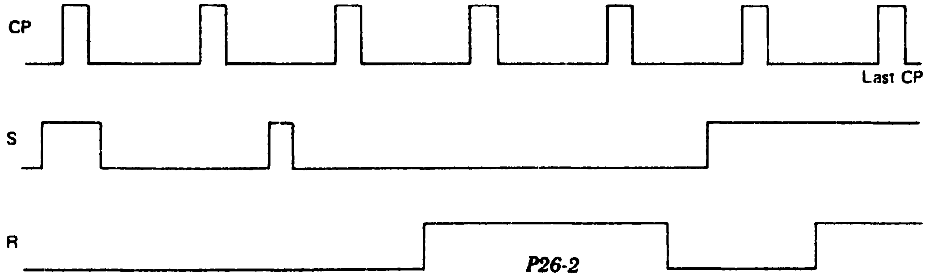
(d) What disadvantage does the DVM in Fig. 26-17 have if V is large and varies during the interval, W?

PROBLEMS

P26-1. An R-S unclocked flip-flop has the waveform shown applied to it. Sketch the corresponding waveforms on the Q and $\bar{Q}$ outputs, assuming the device is constructed from RTL NOR gates.



P26-1



P26-2

P26-2. A clocked R-S flip-flop constructed from NAND gates as in Fig. 26-3(a) has the input waveform shown applied to it. Draw the output waveforms on Q and $\bar{Q}$, assuming change of state takes place on:

(a) the positive rising edge,

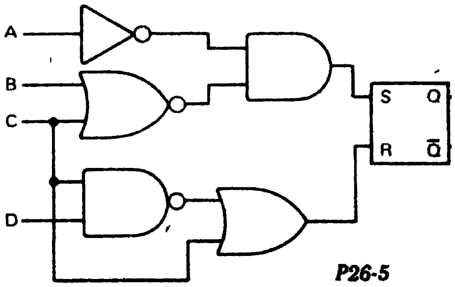
(b) the negative-going edge, of the CP.

P26-3. A CP train consists of 0.5 μsec duration pulses, at a rate of 200 kHz. How long is one bit time?

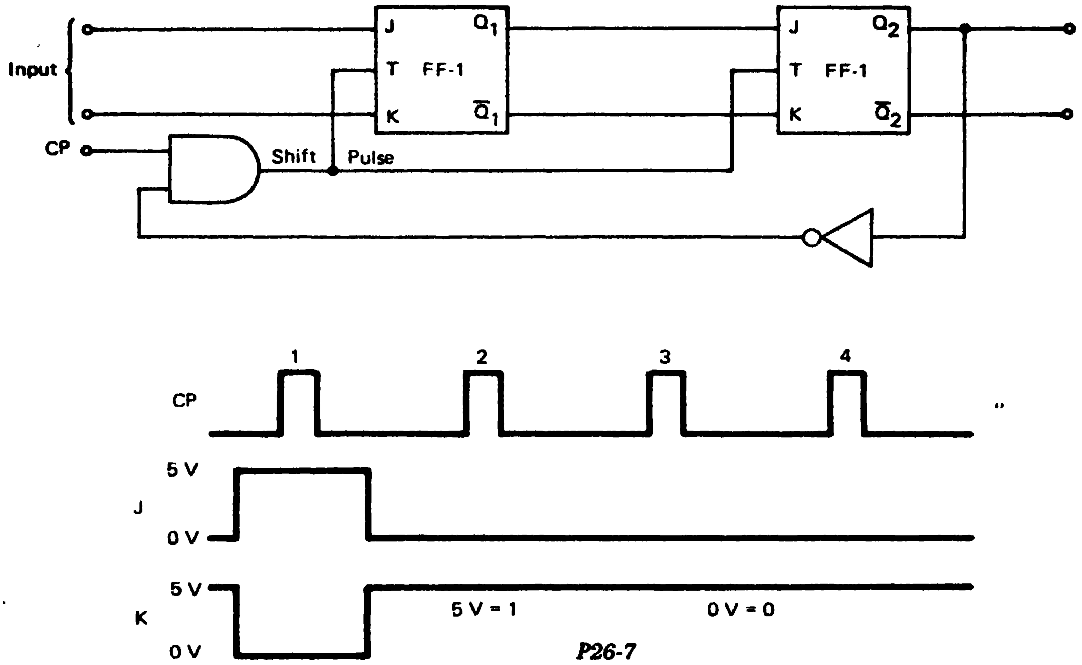
P26-4. If, in P26-2, a direct clear input were applied directly following the last CP shown, what would be the flip-flop states? What if a direct set had been applied?

P26-5. In the logic diagram shown, what would be the Q and $\bar{Q}$ output states of the flip-flop if A = B = C = 1? (Remember $\bar{C}$ + C = 1.)

P26-6. A clocked J-K flip-flop has the waveform shown in P26-2 applied to it. Draw the output waveforms on Q and $\bar{Q}$ if the flip-flop changes state on the negative-going edge of the CP and the flip-flop belongs to the following families: (a) TTL, (b) RTL, (c) DTL.



P26-5


Input
J
Q_1
T FF-1
K
$\overline{Q}_1$
J
Q_2
T FF-1
K
$\overline{Q}_2$
CP
Shift
Pulse
1
2
3
4
CP
5 V
J
0 V
5 V
K
0 V
5 V = 1
0 V = 0

P26-7

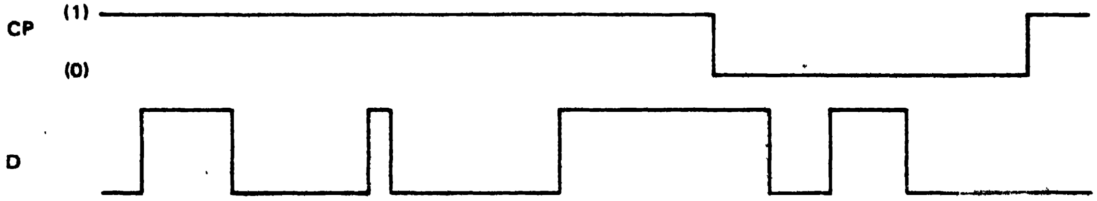
P26-7. Given the two master-slave flip-flops connected as shown and the given input to the J-K terminals, determine the Q and $\overline{Q}$ states of the two flip-flops *after* three clock pulses. Initially, both flip-flops are reset: $Q_1 = Q_2 = 0; \overline{Q}_1 = \overline{Q}_2 = 1$.

P26-8. A D-type flip-flop or bistable latch has the given input applied to its D terminal. Draw the Q and $\overline{Q}$ outputs for the given CP waveform.

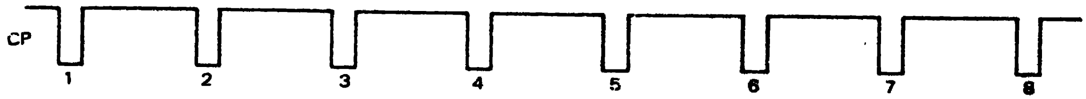
P26-9. A T-type flip-flop function is to be obtained from a J-K flip-flop. What connections must be made at the inputs if the flip-flops belong to the following families:

(a) DTL? (b) RTL? (c) TTL?

P26-10. A T-type flip-flop has the CP shown applied to it. Draw the Q output waveform for the following conditions:



P26-8



P26-10

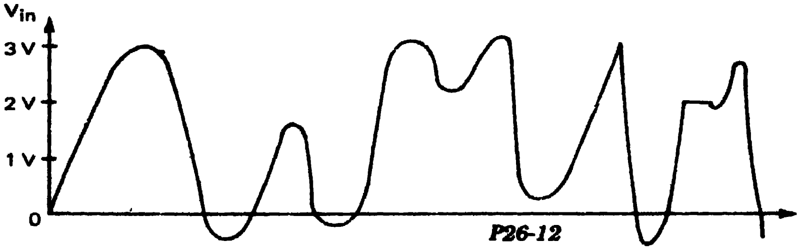
(a) The flip-flop is initially reset, and a momentary direct set is applied between clock pulses 6 and 7.
(b) The flip-flop is initially set, and a momentary direct reset is applied between clock pulses 5 and 6.

P26-11. It is required to produce a pulse of duration 100 msec delayed from a positive input pulse by 20 sec. Specify the R and C values to use with each of the two one-shots shown in Fig. 26-10(a) if the pulse duration of each is given by: T = RC ln 2. (This is the relationship for the IC monostable multivibrator in the SN74121.) The maximum value of R that can be used is 40 K ohms, and the maximum value for C is 1000 μF. Select nearest standard sizes of components.

P26-13. A Schmitt trigger constructed from 2 NOR gates as in Fig. 26-11 has a 2.2-K ohm resistor connected from pin 3 to pin 11, a 470-ohm resistor connected from pin 5 to 11, and a 200-ohm resistor connected from pin 4 to ground. (Ignore any other external resistors shown in Fig. 26-11.) $V_{CC} = 3.6$ V.

Assuming that a saturated transistor has $V_{CE} = 0.2$ V and $V_{BE} = 0.75$ V is necessary for switching, determine:

(a) The quiescent output voltage.
(b) The turn-on threshold voltage.
(c) The unloaded output voltage when in its other stable state.
(d) The turn-off threshold voltage.
(e) The V_{out} versus V_{in} hysteresis curve.



P26-12

P26-12. A Schmitt trigger has an input turn-on threshold voltage of 2 V and a turn-off threshold voltage of 1.5 V. If its quiescent output voltage is 0.8 V and its other stable state provides 5 V, sketch the following:

(a) A graph of V_{out} against V_{in} to show the circuit's hysteresis.
(b) The output waveform for the given input waveform.

P26-14. Repeat P26-13 with $R_e = 100$ ohms.

P26-15. (a) Construct the truth table for a three flip-flop binary counter.
(b) Show how to decode counts of 1, 3 and 7.

P26-16. (a) What is the modulus of a binary ripple counter that employs eight flip-flops?
(b) What maximum frequency division is

this counter capable of? What is the minimum value?

(c) If the input CP has a frequency of 256 kHz, what would be the frequencies at the outputs of the third, fifth, and eighth flip-flops?

P26-17. (a) Construct the waveforms at the outputs of the four flip-flops in a BCD decade counter of the type shown in Fig. 26-14.

(b) If three BCD decade counters are cascaded, what will be the Q states of each of the 12 flip-flops for the following counts:

(i) 57, (ii) 208, (iii) 999.

List the MSB first.

P26-18. A digital frequency counter's variable-gate generator is producing pulses of 0.1 sec duration. The counter indicates 1278. What frequency does this represent?

P26-19. A digital voltmeter is run by a 10 MHz clock. The voltage-controlled generator produces a pulse of duration W = 0.001 V secs, where V is the dc input voltage in volts.

(a) If V = 5 V what will the counter indicate?

(b) What voltage does a count of 5370 represent?

(c) How many cascaded BCD decade counters are needed if four significant figures are to be displayed?

P26-20. (a) Refer to Fig. 26-19. Assume flip-flop 3 has just been set high, and G-6 sets flip-flop 2 high. What will happen?

(b) What is the maximum length of time the left-turn traffic will be allowed to move for the condition in (a)? What will terminate it?

(c) If flip-flop 2 has not been set, but flip-flop 1 has, how long must the first (and only) car wait?

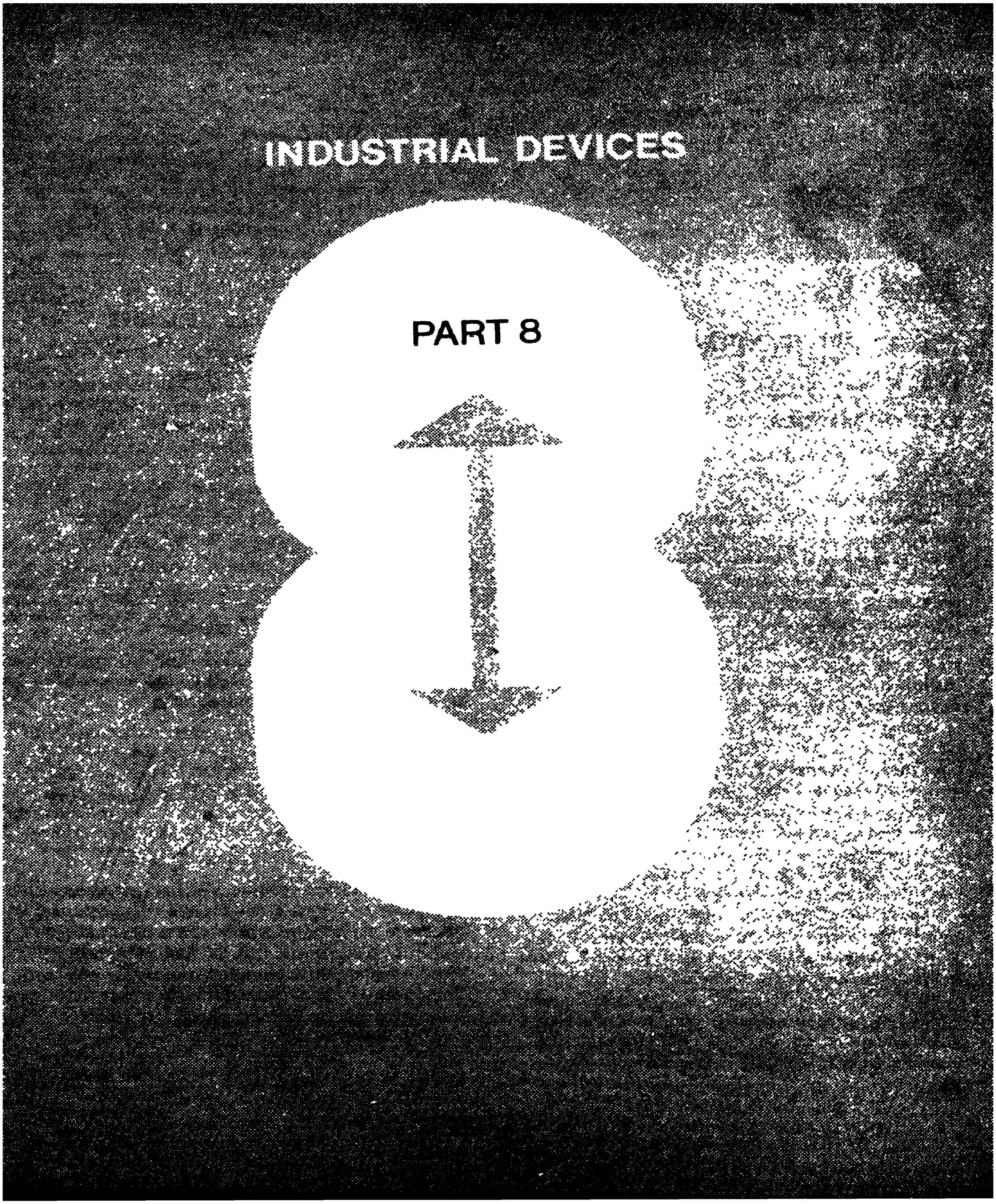
(d) If, after this car starts to move, more cars continue to arrive to make the left turn without letting the output of G-8 go high, how long a time will they be permitted to keep the light green?

(e) If a long line of left-turn traffic is waiting and only the first two or three cars can make the turn before the light turns back to red, what possibly could be the problem in the one-shot settings?

(f) In some cities, a right turn on a red light is permitted if the main artery is clear in that direction. If a solitary approaching car activates sensor 1 but then moves into a right-turn lane that does not include sensor 2, where would you place a third loop (or more) to sense if this car makes a successful right turn? How would you use this signal to deactivate the sequence leading to a left-turn green signal caused by this car?

INDUSTRIAL DEVICES

PART 8



27

OPTOELECTRONIC DEVICES

27-0. INTRODUCTION

The relation between light and electricity was first observed by Hertz in 1887 and was termed the photoelectric effect.

Although this effect includes the generation of electron-hole pairs in semiconductors, it initially referred to the emission of electrons from a metallic surface. The devices that resulted from this emission were vacuum and gas phototubes and were commonly called photoelectric devices. However, modern solid-state devices, which include emitters, sensors and couplers (see Fig. 27-1), usually are known as optical electronic devices or optoelectronics for short. Electro-optics is another name sometimes used.

As can be seen from Fig. 27-1, there are two other categories in the area of optoelectronic sensors besides the photoemissive type. The photovoltaic devices generate a voltage, while the photoconductive devices, which require an external source, may be further subdivided into two branches. One is the bulk type often referred to as photoresistors; the other uses reverse-biased depletion junctions and includes photodiodes and phototransistors.

Optoelectronic solid-state emitters are of three main types. There are light-emitting diodes (LEDs) that emit visible radiation for display purposes, infrared (IR-type) emitters, and laser diodes.

Finally, emitters and sensors are often combined in one package to provide a coupling function, such as providing a high degree of isolation between low-powered integrated circuits and ac power line volts, etc.

Since the characteristics of optoelectronic devices involve some terminology from the field of light measurement, we should briefly review the basic fundamentals of light that will be pertinent to our discussion.

27-1. FUNDAMENTALS OF LIGHT

The Electromagnetic Spectrum

What we commonly call light occupies only a small portion of the electromagnetic spectrum.

This is a graphical representation showing the different radiations that exist at different frequencies or wavelengths. Fig. 27-2 shows a portion of the electromagnetic spectrum where the wavelength is given in millimicrons or nanometers.

1 micron = 1×10^{-6} meter = $1\ \mu$.

1 millimicron = 1×10^{-9} meter = $1 \mathrm{m}\mu$.

1 nanometer = 1×10^{-9} meter = 1 nm.

The frequency and wavelength of electromagnetic energy are related to the speed of light by the following equation:

$$c = f\lambda, \qquad (27\text{-}1)$$

where

c = velocity of light = 3×10^{8} m/sec,

f = frequency in Hz,

λ = wavelength in meters.

Another characteristic of light or radiant energy was suggested by Planck's hypothesis. He made the assumption that radiant energy was not continuous, but existed in discrete bundles or quantities called *quanta* or photons. The energy of a photon depends upon the frequency of the radiant energy as given by the following equation:

$$E = hf, \qquad (27\text{-}2)$$

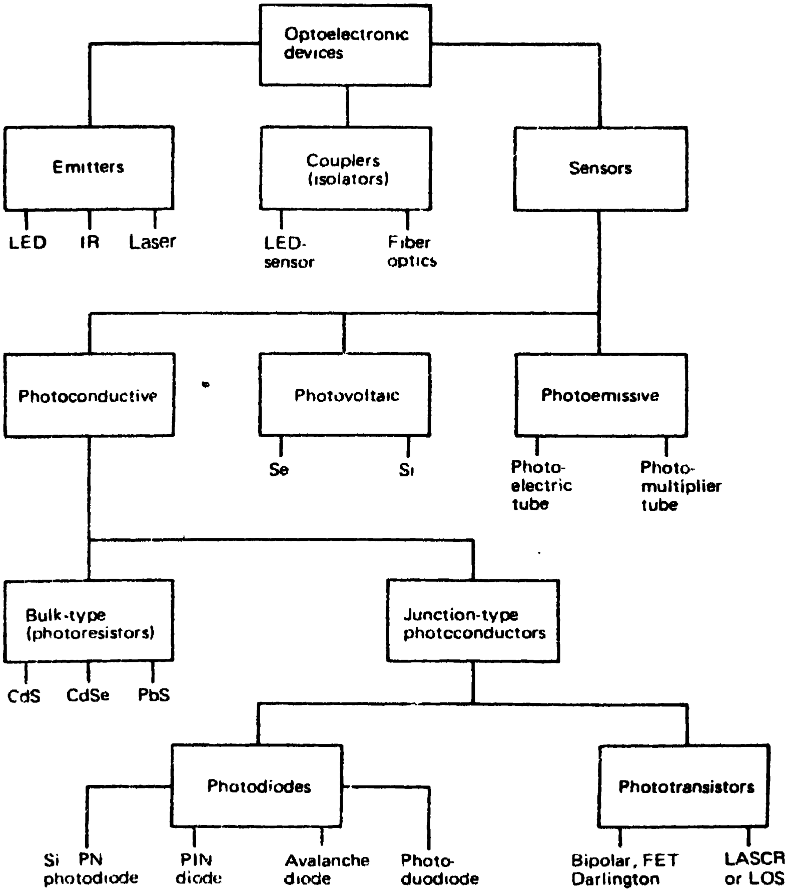
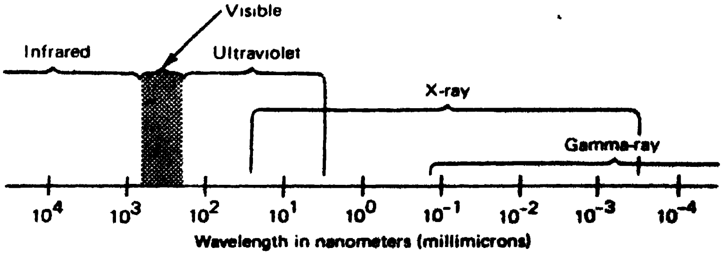

Optoelectronic devices
Emitters
Couplers (isolators)
Sensors
LED
IR
Laser
LED-sensor
Fiber optics
Photoconductive
Photovoltaic
Photoemissive
Se
Si
Photo-electric tube
Photo-multiplier tube
Bulk-type (photoresistors)
Junction-type photoconductors
CdS
CdSe
PbS
Photodiodes
Phototransistors
Si PN photodiode
PIN diode
Avalanche diode
Photo-duodiode
Bipolar, FET Darlington
LASCR or LOS

Fig. 27-1. Showing the devices that make up the field of optoelectronics.

Fig. 27-2. Portion of the electromagnetic spectrum.


Visible
Infrared
Ultraviolet
X-ray
Gamma-ray
10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
Wavelength in nanometers (millimicrons)

where

E = the photon energy in joules,

f = the frequency in Hz,

h = Planck's constant = 6.626×10^{-34} joule-sec.

Thus, at higher frequencies (smaller wavelengths), the energy of the photons is higher and can become very penetrating. This explains why X-rays and gamma-rays (see Fig. 27-2) must be handled very carefully. We will use the concept of photons to later explain the photoconductive effect.

Spectral Response

Optoelectronic devices utilize energy in the visible and infrared regions. Fig. 27-3(a) shows the response of the average human eye to the visible radiation that extends from 0.4 to 0.76 microns in wavelength. The eye has a maximum response to light at 0.55 μ and drops to 50% of maximum at 0.61 μ.

The tungsten incandescent lamp is often used as a source of radiation in photoconductive devices. As Fig. 27-3(b) shows, its peak radiation energy output is far beyond the visible spectrum into the infrared, and it may adequately be used as a source of infrared radiation in many applications.

The curve drawn for the tungsten lamp in Fig. 27-3(b) is at its standard operating temperature and corresponds to a color temperature of 2870°K.* For lower color temperatures the peak moves more to the right, with very little visible radiation at 1600°K. The sun's radiation (color temperature 6000°K) has its peak in the middle of the visible region, but its energy extends deep into the infrared.

Note that the wavelength in Fig. 27-3(b) is also given in angstroms, Å, a quite often used unit in optics.

1 angstrom = 1Å = 1×10^{-10}meter.

10 Å = 1 mμ.

Illumination and Irradiance

(a) Illumination

A full discussion leading to the true definition of illumination is beyond the scope of this text, but the following will reveal the basic terminology involved.

To activate a photoconductive device requires a certain radiation or illumination level. These are two different methods of specifying the level involved and are measured in two different types of unit. An *illumination* level is generally used where a device is primarily responsive to visible light. In this case, the unit used is the *footcandle* or lumen/sq ft. A *radiation* level is generally used for devices responsive to energy primarily outside of the visible region. Then the unit of measurement used is the mW/cm^2. However, the unit of measurement used for a device will vary from one manufacturer to another with some supplying information both in footcandles and mW/cm^2.

At one time, the standard source of illumination was the light given off by the international or standard candle, in which a 7/8 in. diameter spermaceti candle, burned at a rate of 7.776 gms/hr.* This originated the term

---

*The color temperature of a source is the temperature at which an ideal blackbody radiator must be operated to produce the same spectral distribution or visual effect as the source in question.

*The modern standard source consists of fused thorium oxide immersed in a chamber of platinum. When the platinum is at its melting point, the light emitted from the chamber and contained in one steradian (a unit solid angle equivalent to a three dimensional radian = A/r^2, $A = r^2$) is said to be a *lumen* of luminous flux.

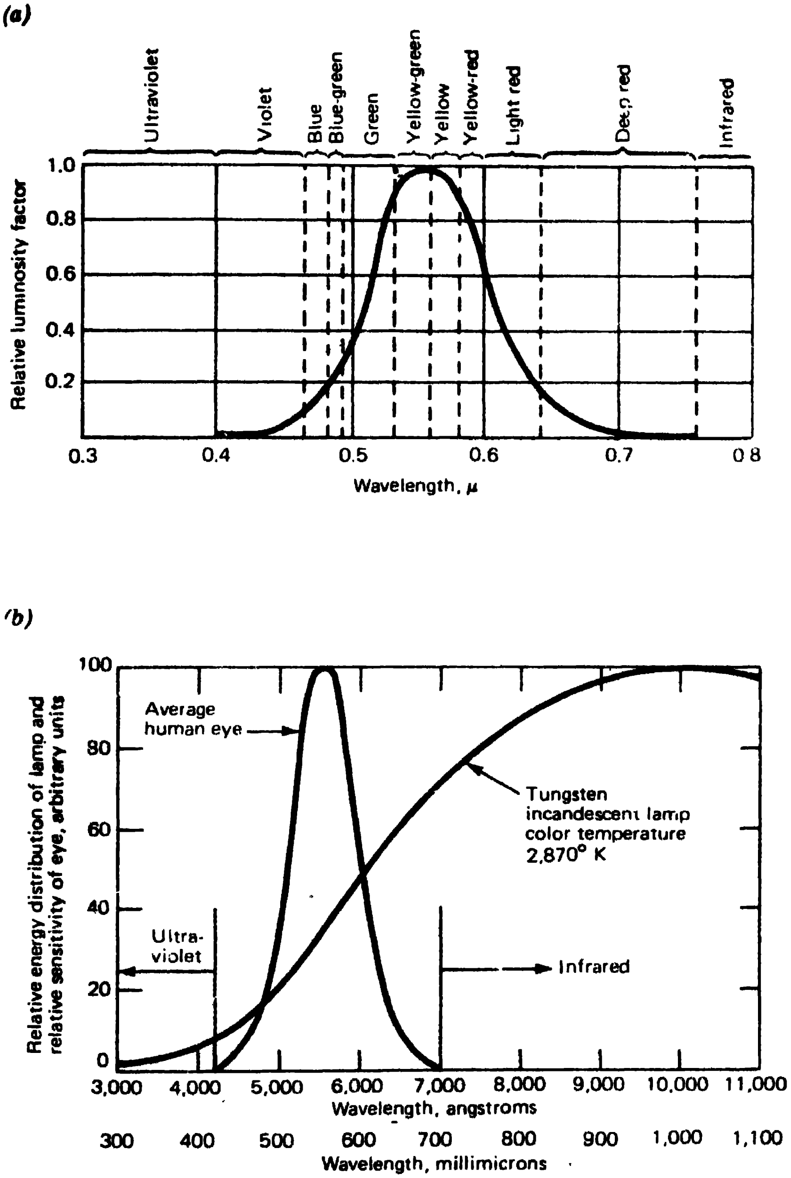


Fig. 27-3. (a) Spectral response of the eye to visible radiation. (b) Comparison of the spectral response of the eye and the radiation characteristic of a tungsten lamp working at a standard color temperature.

candlepower, which is still used as a measure of the *luminous flux* emitted from a light source. A surface is said to have an *illuminance* (usually but incorrectly referred to as an *illumination*) of 1 footcandle when the surface is placed 1 foot away from a 1-candlepower source.

A more modern unit of luminous flux is the lumen, which is defined so that 4π lumens = 1 candlepower.

If the illuminance is constant over the surface area,

$$E = F/A, \tag{27-3}$$

where

E = the illuminance or luminous flux density in lumens/sq ft,

F = luminous flux in lumens,

A = area in sq ft.

If we consider a point source of 1 candlepower or 4π lumens and a sphere of 1 foot radius around it, what will be the illuminance on the sphere?

$$E = F/A = \frac{4\pi \text{ lumens}}{4\pi \text{ sq ft}} = 1 \text{ lumen/sq ft.}$$

That is, the surface area of the sphere is $4\pi r^2 = 4\pi$ sq ft.

But a 1 candlepower source 1 foot distant from a surface is 1 footcandle. Hence, 1 lumen/sq ft = 1 footcandle.

The illuminance at a point may be measured using a photometer, which is a device to compare the illuminance at the point with that from a standard source. Alternatively, a footcandle meter which utilizes the photoelectric effect may be used to indicate directly the illuminance. The deflection of a meter responding to current flowing through a photoconductive cell is calibrated to read directly in footcandles. As a standard of comparison, the illuminance from direct sunlight is in the order of 10,000 footcandles, twilight provides 1 footcandle, and a full moon 0.01 footcandle. A well lighted surface for reading should have from 50 to 100 footcandles illuminance (illumination).

(b) Irradiance

Illuminance is an indication of the *visible* radiation given off by a source. For example, a 100-watt tungsten lamp typically emits a flux of 1630 lumens, while a 100-watt 60-inch fluorescent lamp emits 4,400 lumens. However, if a photoconductive cell, responsive to energy in the visible and in the infrared range, is exposed to a tungsten source the illuminance in footcandles is not a proper indication of the radiant energy on the cell. We thus need to define the radiation flux density on a surface, the irradiance, in terms of the *total power* from the source in *watts*. This will include both the visible and invisible light energy.

If the irradiance is constant over the surface area,

$$H = P/A, \tag{27-4}$$

where

H = irradiance or radiant flux density in watts/cm^2,

P = total radiated energy in watts,

A = area in sq cm.

The determination of total *radiated* energy in watts is not very easy (it may be done using a solar cell), and no direct conversion between lumens and watts is possible (without involved calculations) because of the variation of lumens with frequency.* Typical values of irradiance on photoconductive devices range from 1 to 20 mW/cm^2. For comparison, the sun delivers approximately 100 mW/cm^2 on an average sunny day, and a 2-cell flashlight approximately 10 mW/cm^2. A tungsten filament lamp source operating at 2870° K and producing an illuminance of 20 footcandles would correspond to an irradiance of approximately 1 mW/cm^2.

---

*1 watt of radiative power is equivalent to 680 lumens at a frequency of 0.555 μ, or 1 mW/cm^2 = 1.58×10^{-3} lumens/ft^2.

27-2. PHOTOCONDUCTIVE SENSORS

Bulk-Type Photoconductive Cells

The increase in conductivity of a semiconductor when exposed to light may be explained by referring to Fig. 27-4.

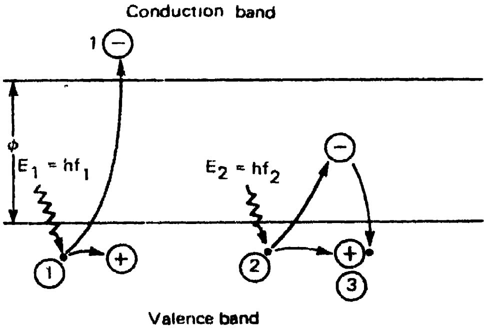


Fig. 27-4. Showing the photoeffect in a semiconductor. The semiconductor's crystal energy structure is represented by the valence and conduction bands.

Consider a photon striking the surface of a semiconductor. If its energy, $E_1 = hf_1$, is greater than the energy gap between the valence and conduction energy bands of the semiconductor, sufficient energy will be imparted to an electron to raise it to the conduction level, leaving a hole in the valence band. This electron-hole pair is free to serve as current carriers, so that conductivity of the material increases, with the increase depending upon the illumination and area, as shown in Fig. 27-5.

Since these bulk-type photoconductive cells are characterised by a change in resistance, they are often referred to as photoresistors. They are specified in terms of the change in resistance that takes place from 2 footcandles to dark within a 5-second interval. For example, a Clairex CL905L, (curve C in Fig. 27-5) has a resistance of 10 K ohms at 2 footcandles and a minimum resistance ratio of 1:100, indicating an ultimate resistance of at least 1 M ohm in the dark (0.01 footcandle). Encased, the CL905L is 0.21 inch in diameter and 0.15 inch long.

The spectral response of photoconductive cells depends upon the basic material used, cadmium sulfide (CdS), cadmium selenide (CdSe), or lead sulfide (PbS), as well as the doping and treatment used. See Fig. 27-6.

The fact that the curves show a peak response at different wavelengths indicates a threshold effect which is different for each material. Photons having insufficient energy to excite electrons to the conduction band, such as $E_2 = hf_2$ in Fig. 27-4, will not increase conductivity. Cadmium sulfide cells have their peak response in the visible spectrum and by suitable selection can be used in color sorting circuits. A recently announced sensor, a Clairex 7UV 20-S with a sapphire window, has its peak in the ultraviolet region and responds to radiation from 2400 Å to 3900 Å.

The primary advantage of bulk-type photoconductors is their high dissipation capability (up to 2 W with a heat sink) and high voltage capability (100 to 300 V), making it possible to directly control relays and small motors. Fig. 27-7(a) and (b) show two such applications, while in (c) a transistor is used to amplify the resistance change.

The main drawback of bulk-type photoconductive cells is their speed of response. Rise times vary from 1 to 100 msec and fall times from 0.7 to 30 msec, with PbS the fastest and CdS the slowest. Accordingly, their frequency response is limited to a frequency range of

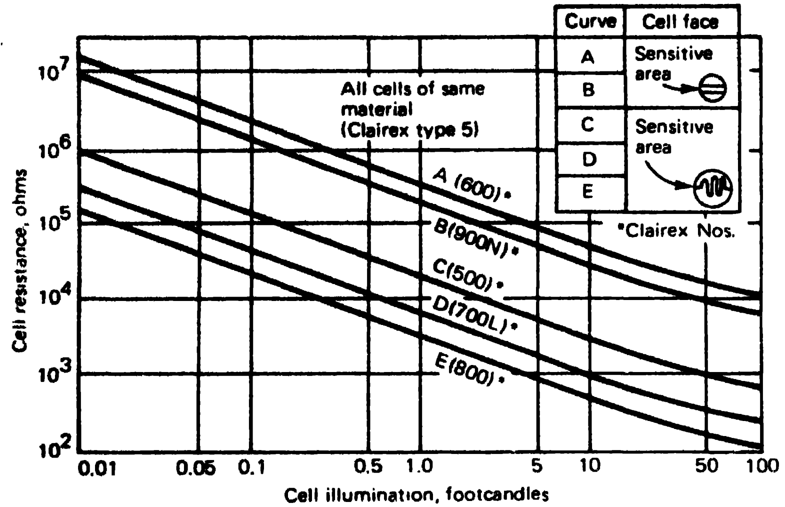


Fig. 27-5. Curves showing how the resistance of a photoconductive cell varies with the illumination and the sensitiv a. For example, cell B is larger in area than A but has the same pattern. Cells C, D, and E have a wavy pattern ulting in a larger area. All cells are CdS. (Courtesy of Clairex Electronics.)

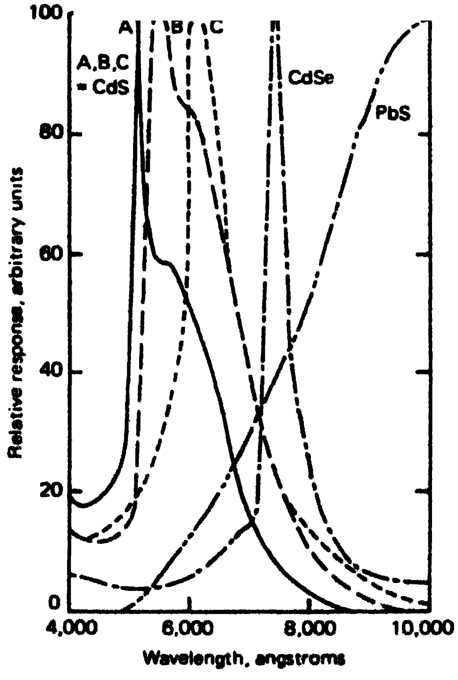


Fig. 27-6. Curves showing the spectral response of different photoconductive cells and materials. Curves A, B, and C are all cadmium sulfide with different doping and treatment. (Courtesy of Clairex Electronics.)

from 100 to 5000 Hz. Nevertheless, they are very economical and are widely used in low-speed switching, counting, and beam-breaking applications, as well as in cameras to control shutter opening during the flash. Twin photoconductive cells mounted in the same package have been used in optical bridge circuits for position control mechanisms and dual-channel remote volume control circuits.

Junction-Type Photoconductors

(a) PN Photodiodes

Bulk-type photoconductive cells could also be fabricated from silicon or germanium, but the current sensitivity to radiation can be made much larger by the use of a reverse-biased PN junction, as shown in Fig. 27-8(a).

The mechanism of current control through radiation is similar to that explained for the photoconductive cell, whereby photons create electron-hole pairs on both sides of the junction. With no light applied, the reverse current is simply the reverse saturation current due to

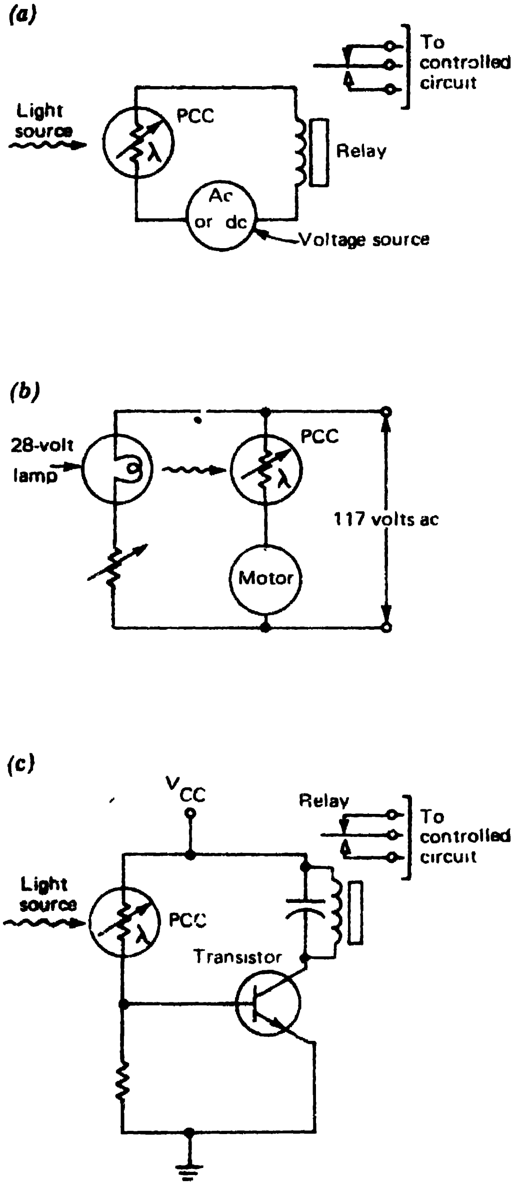


Fig. 27-7. (a) A photoconductive cell operates a relay directly. (b) A variable resistance changes the light level from a lamp focused on a power photoconductive cell which directly controls the speed of a small appliance motor. (c) A transistor amplifies the resistance change in the cell to operate the relay.

the minority carriers, holes in the N-type and electrons in the P-type. When light is applied, the photoinduced electrons in the conduction band of the P-type will move across the junction down the potential hill with the thermally generated minority carriers. Similarly, holes produced in the valence band of the N-type (by electrons being photoexcited into the conduction band) are available to add to the current flow by moving across the junction to the P side. The spectral response of the diode to wavelength is shown in Fig. 27-8(b).

As shown by Fig. 27-9, the photoinduced current or light current is a function of the irradiance. With $H = 0$, the dark current, I_D, is typically 30 to 70 nA at $V_R = 30$ V. When $H = 20$ mW/cm^2, the light current, I_L, is typically 10 to 90 μA at $V_R = 5$ V.

The active diameter of these devices is only 0.1 inch, but they are mounted in standard TO-5 packages with a window. They can operate at frequencies in the order of 1 MHz and are used in character recognition, optical communications equipment, and encoders. Since the application circuits involving photodiodes are similar to those for phototransistors, we will examine practical applications later.

(b) PIN Photodiodes

Although the PN photodiode provided a much higher frequency response than the CdS bulk type photoconductors, an even faster response may be achieved in the PIN photodiode. This consists of adding a layer of *I*ntrinsic silicon between the heavily doped *P* and *N* type silicon materials. This has the effect of reducing the transit or diffusion time of photoinduced electron-hole pairs. That is, carriers created by photons incident on the *middle* of the depletion region have less distance to travel than if generated at one side or the other of the depletion layer. The relatively thick layer (2.5 μ) of high resistivity intrinsic silicon

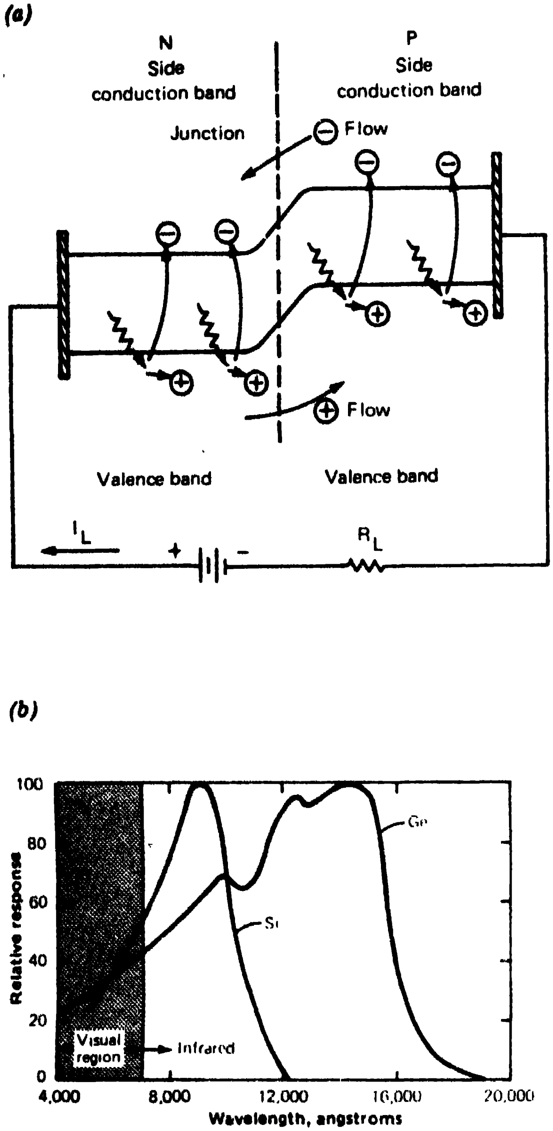


Fig. 27-8. (a) Showing the photoeffect in a reverse-biased PN junction. (b) Curves showing the spectral response of Si and Ge photosensors compared with the visible region.

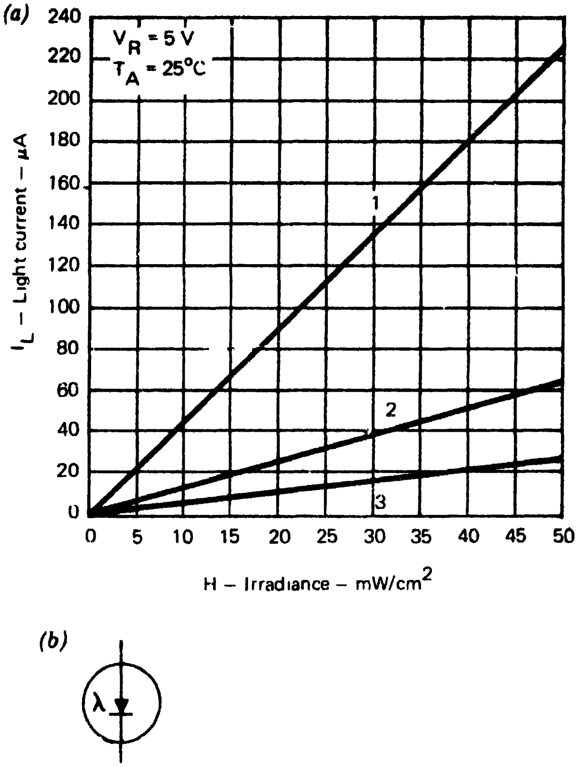


Fig. 27-9. (a) Showing how light current is linearly related to irradiance for three different silicon PN epitaxial planar photodiodes. Source for H is an unfiltered tungsten filament bulb operating at a color temperature of 2870°K. (b) Symbol of a PN photodiode.

ensures the absorption of most of the incident photons.

The electrical (and optical) characteristics of the MRD 500 PIN photodiode are indicated in Fig. 27-10. Note the linearity of light current with irradiance and how the device is essentially a constant-current generator at a given light level.

The device has a typical radiation sensitivity (from a tungsten source) $S_R = 1.8\ \mu A/mW/cm^2$, at $V_R = 20$ V. This compares with an S_R of

0.4 to 4 μA/mW/cm^2 for the silicon PN photodiodes. However, the response time of the PIN photodiode is ultrafast, with a switching speed of typically 1 nsec. It is thus used to detect laser pulses, visible and infrared light-emitting diodes, and in switching and logic circuits.

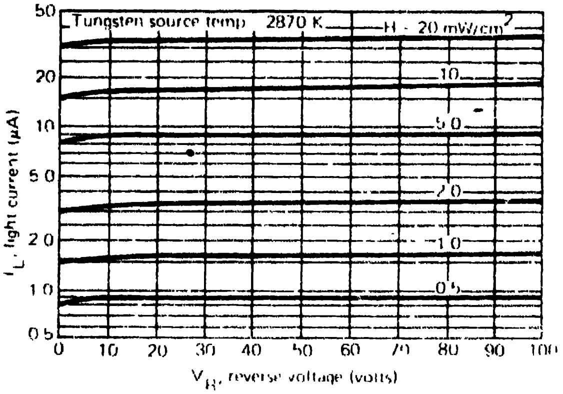


Fig. 27-10. Irradiated voltage-current characteristics for the MRD 500 PIN photodiode. (Courtesy of Motorola Inc.).

(c) Avalanche Photodiode

Current sensitivity may be increased from 30 to 100 times by operating a photodiode in its reverse breakdown or avalanche breakdown region. In the avalanche photodiode, the construction is such as to provide a very uniform junction that exhibits the avalanche effect at voltages between 30 and 200 V for the TIXL55-59 series. The dark current is typically 10 μA. The electron-hole pairs that are generated by incident photons are accelerated by the high electric field to "kick" new electrons from the valence to the conduction band, as explained in Section 5-14. In this way, a typical photomultiplication of 50 is obtained, so that light current may be 0.5 to 1 mA.

It is to be noted that this mechanism does not reduce the fast response found in the PIN photodiode. The avalanche photodiode may be operated at a modulation frequency of 2 GHz with an excellent signal-to-noise ratio that rivals the photomultiplier tube. Their present disadvantage is their expense.

(d) NPN Photoduodiode

The current flow in the photodiode thus far has been relatively small. In an attempt to obtain an *inexpensive* device with higher current capabilities, another junction was added to form a duodiode. Resembling a transistor but differing in the 10 times wider base, a photoduodiode such as the 1N2175 has a typical sensitivity of 22 μA/mW/cm^2, roughly 10 times that of the previous devices. It has a typical dark current of 0.01 μA, a rise time of 2 μsec and fall time of 45 μsec, and a power dissipation of 250 mW.

The 1N2175 is only 0.082 inch in diameter and 0.4 inch long. It is thus well suited for punched-card and tape readout applications requiring close mounting centers. Because of its two symmetrical PN junctions, it can be used with either ac or dc bias voltages. Symbols used for NPN photoduodiodes are shown in Fig. 27-11.

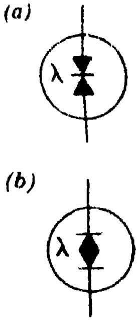


Fig. 27-11. Symbols used for NPN photoduodiodes.
(a) Usual symbol.
(b) Alternate older symbol.

(e) NPN Phototransistors

The next logical step to increase radiation sensitivity was to use the inherent current multiplication found in a transistor. Actually, any bipolar transistor is affected by light on its collector-base junction, and for this reason transistors are encased in light-tight packages. However, phototransistors are constructed to optimize their light-activating characteristics, and often have a lens built into the package to focus radiation on the collector-base junction. Then the photoinduced current of this junction (like that in a photodiode), I_λ, serves as the base current of the transistor. The amplification or current gain of the transistor then results in a collector current of

$$I_C = (h_{fe} + 1)I_\lambda. \qquad (27\text{-}5)$$

Fig. 27-12(a) shows that the transistor is connected in a conventional manner, so that the collector-base junction is reverse biased. The base lead may be provided in some packages and may be left floating or used to bias the transistor into some area of operation.

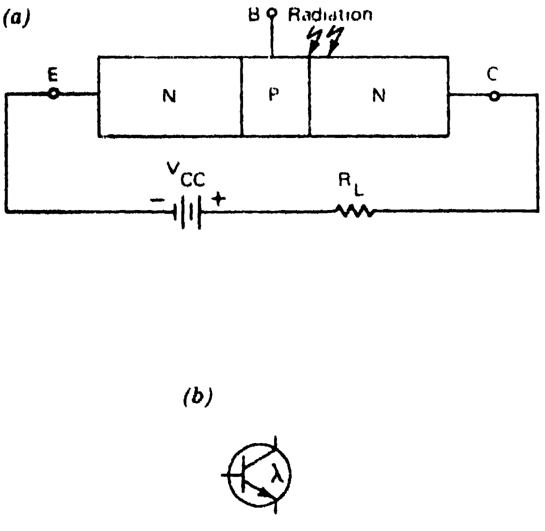


Fig. 27-12. (a) Sketch of a phototransistor showing how radiation is directed onto the collector-base junction which is reverse-biased by V_{CC}. (b) Symbol of the NPN phototransistor.

The electrical characteristics of an FPT120 high-sensitivity phototransistor are shown in Fig. 27-13(b). Note how they resemble the CE characteristics with radiant energy the controlling parameter. A sensitivity S = 2.5 mA/mW/cm^2 is evident for this device, and the curve in Fig. 27-13(a) shows how this may be related to illumination in footcandles. A partial price paid for this high sensitivity is reduced speed of operation, although t_r and t_f for the light current are still only 18 μsec with a 100-ohm load.

The device, with lens, is typically 0.2 inch in diameter and 0.2 inch long, not including the leads. It has an ambient dissipation capability of 100 mW, a dark current of typically 10 nA, and a maximum collector current of 25 mA. It may be used like a photodiode in applications requiring high sensitivity.

(f) Other Phototransistor Types

The sensitivity of the FPT120 phototransistor described above is unusually high. More typical values for other types run from 0.02 to 0.8 mA/mW/cm^2. Another method of obtaining sensitivities in the range of 1 to 4 mA/mW/cm^2 -- and more economically — is to use a *photo Darlington amplifier.* This contains a silicon phototransistor and a direct-coupled transistor amplifier stage. See the symbol in Fig. 27-14. The disadvantage is a much higher switching time, typically 250 to 400 μsec.

The photo-FET is relatively a newcomer to the field of optoelectronics, but it has a high power gain and fast response. This is especially true when cooled to -198°C (in liquid nitrogen), where its g_m increases to 3 times its room temperature value, and its rise time is only 30 nsec.

Although not strictly a transistor, the photo-SCR (silicon controlled rectifier) belongs to

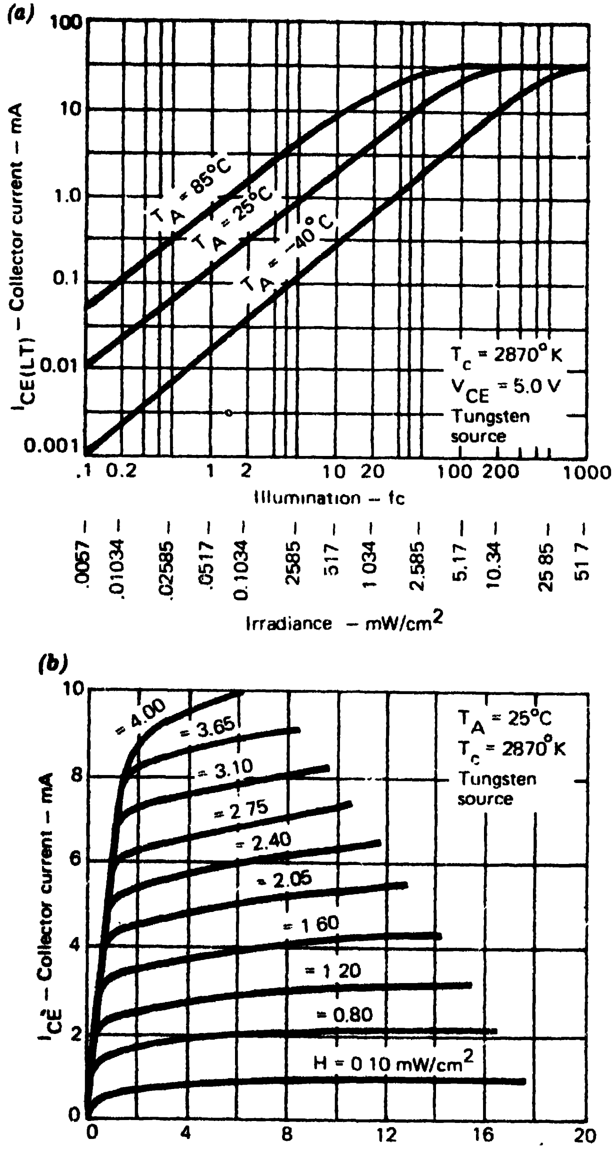


Fig. 27-13. (a) Photo current characteristics of the FPT120 NPN phototransistor showing the variation of collector current with temperature and illumination. (b) The collector current versus collector voltage characteristics of the FPT120 NPN phototransistor. (Courtesy of Fairchild Semiconductor.)

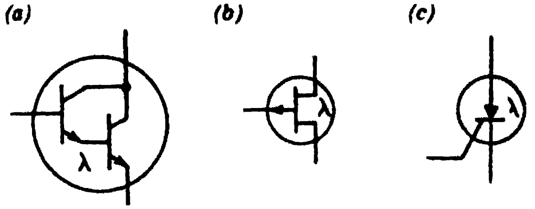


Fig. 27-14. (a) Symbol of a photo Darlington amplifier. (b) Symbol of a photo-FET. (c) Symbol of a photo-SCR or LASCR.

the junction-type photoconductive family. A pulse of light can be used to switch on the current (up to 0.7A at 200 V) through the device but then loses control of the current. A full description of the SCR is given in the next chapter.

27-3. APPLICATIONS OF PHOTODIODES AND PHOTOTRANSISTORS

Light-Operated Relays

It has been shown that junction-type photoconductors have relatively small dissipation capabilities compared with the bulk type. Consequently, they usually require some stages of amplification to provide a control function and are often found in the type of circuit shown in Fig. 27-7(c). Either a photodiode or phototransistor could take the place of the cell, providing actuation of the relay upon illumination of the junction.

Conversely, if the photosensor and resistor R_1 were interchanged, the relay would be actuated upon the interruption of the light beam and could be used to sound an alarm, advance a mechanical counter, or provide some other function. If used as an intrusion alarm, the circuit could be made more foolproof by the use of infrared radiation, which silicon photodiodes and phototransistors respond to readily. If a tungsten source were used, a suitable filter could block the visible portion and permit the passage of the infrared beam to irradiate the sensor.

Both of these applications involve the operation of the photosensor as a switch. The amount of light necessary to make it go into saturation could be determined by the drawing of a load line on the characteristic curves, as with a conventional transistor, with the slope of the load line equal to $1/R_1$.

Electro-optical Shaft Encoder

Another application that uses the phototransistor as a switch is shown in Fig. 27-15. This is an electro-optical shaft encoder that gives a 4-bit binary representation of the angular position of a motor shaft or other rotating body. The heart of the system is the code disc which consists of 16 different patterns of alternate transparent and opaque sections that rotate with the shaft. In any given position, light sources on one side of the disc will illuminate some combination of phototransistors mounted in line on the other side. In the example shown, a binary 0101 will be registered, indicating the shaft is in position "5" of 16 possible positions. If this binary number is compared with the desired position, say 1101, the shaft will continue to rotate until the output of a comparator circuit shows equivalence. With 16 positions (2^4) the shaft will be placed to a desired position with a maximum error of ±11.25°. Increasing the number of bits by using a monolithic linear array of 16 phototransistors would give a maximum positional error of only ±0.165 minutes of arc, and the system would more than likely be limited in accuracy by the motor and gear train.

Such an optical shaft encoder finds applications in fire-control systems, radar-tracking systems, multiaxis numerically controlled machines, etc.

Optical Paper Tape Reader

With minor modifications, the above system (without the code wheel) can be used to "read" punched cards and tapes by sensing the holes. One track of the tape or cards has a continuous line of holes that provides a clock pulse function; this CP input synchronizes the logic circuitry with the pulses arriving from the phototransistors on each track or channel.

Brushless Dc Motor

Another application of phototransistors that is similar to the optical shaft encoder system is in the brushless dc motor. Dc motors are popular for their easy speed control, but the brushes (and commutators) have a limited life, and the arcing that occurs limits the adaptability to spacecraft applications and other hazardous environments.

You will recall that the brushes are used to channel the current to the moving rotor. And the commutator bars direct the current to the correct rotor coils at the proper instant to provide torque. In the brushless dc motor, a permanent magnet is used for the rotor, and three windings are used for the stator. As shown by Fig. 27-16(a), the position of the rotor is sensed by a disc or light shield attached to the rotor and moving between three pairs of light emitters and photosensors. The photosensors drive switches to energize in sequence the stator coils A, B, and C, providing a rotating magnetic field and continuous rotation of the rotor. Improved efficiency and a smaller motor are obtained by using a delta winding with six switches. Its performance compares very favorably with conventional dc brush-type motors. Motors up to 0.5 hp are available at voltages from 10 to 27 V dc.

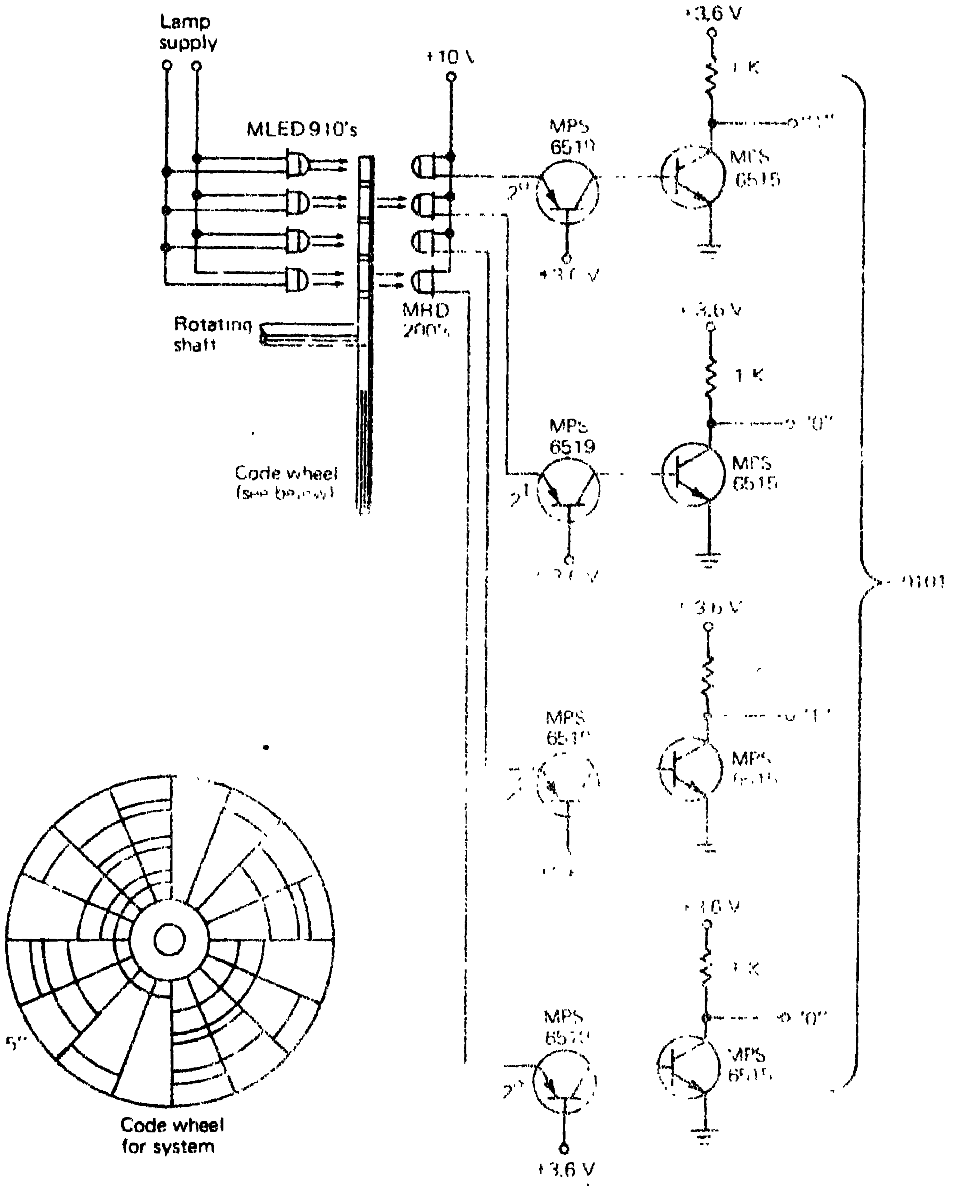

Lamp supply
MLED 910's
+10 V
MRD 200's
Rotating shaft
Code wheel (see below)
+3.6 V
1 K
"1"
MPS 6519
MPS 6515
2^0
2^1
2^2
2^3
"0"
"1"
"0"
0101
5"
Code wheel for system

Fig. 27-15. Electro-optical shaft encoder system showing the code wheel and position "5. (Courtesy of Motorola Semiconductor Products, Inc.)

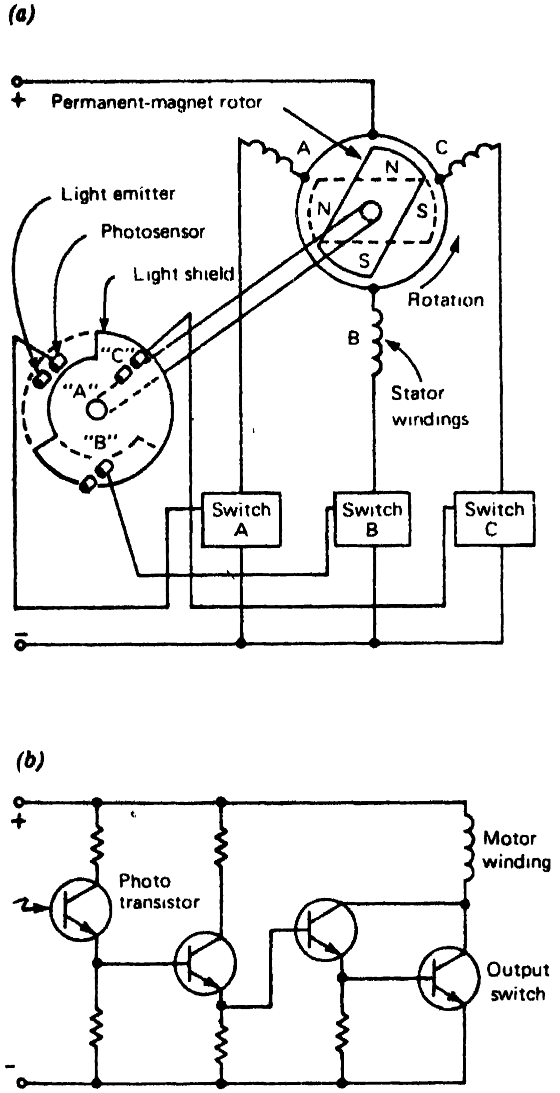

(a)
+
Permanent-magnet rotor
Light emitter
Photosensor
Light shield
A
C
N
S
N
S
"C"
"A"
"B"
Rotation
B
Stator windings
Switch A
Switch B
Switch C
−
(b)
+
Photo transistor
Motor winding
Output switch
−

Fig. 27-16. (a) Schematic diagram of a brushless dc motor (BDCM) showing the light emitter-photosensor electronic commutator.
(b) Typical switching circuit used in the electronic commutator. (Courtesy of Control Engineering.)

Optoelectronic Spring-Mass Accelerometer*

Not all photosensors are used as switches. Fig. 27-17(a) shows an accelerometer in which the amount of light passing from the source to the phototransistor sensor is determined by the position of a shutter blade. The position of the spring blade varies, due to its mass, as a force acts upon it when the accelerometer is undergoing acceleration in a direction perpendicular to its base. The output voltage is directly proportional to the acceleration and may be calibrated to give a direct readout. Typical sensitivity is 1 volt per "g" of acceleration.

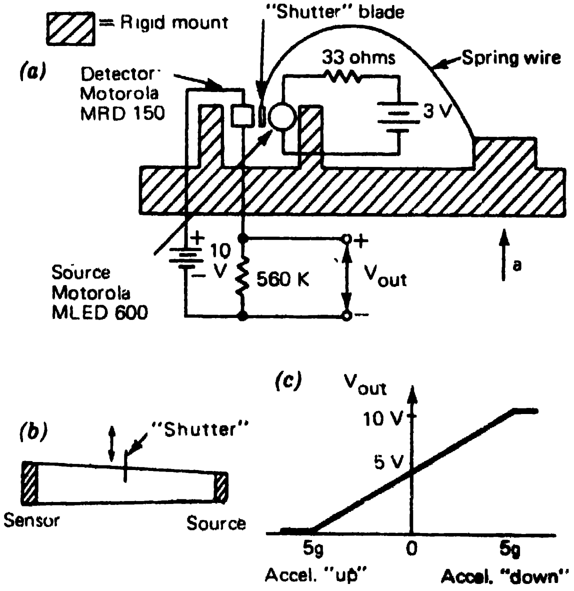

▨ = Rigid mount
"Shutter" blade
(a)
Detector Motorola MRD 150
33 ohms
Spring wire
3 V
Source Motorola MLED 600
10 V
560 K
+
Vout
−
a
(b)
"Shutter"
Sensor
Source
(c)
Vout
10 V
5 V
5g
0
5g
Accel. "up"
Accel. "down"

Fig. 27-17. (a) Schematic diagram of an optoelectronic accelerometer. (b) Showing how the shutter controls the light beam (shown for a = 0). (c) Curve showing the typical output voltage as a function of acceleration.

---

*See *The Instrumented Laboratory*, a report developed by the Commission on College Physics, University of Maryland, May 1972.

27-4. PHOTOVOLTAIC SENSORS

If the characteristic curves of the photodiode in Fig. 27-10 are examined, it is found that the curves do not pass through the origin but are shown terminating on the current axis. If the curves were continued to the "left" for *forward voltages*, they would have the form shown in Fig. 27-18(a).

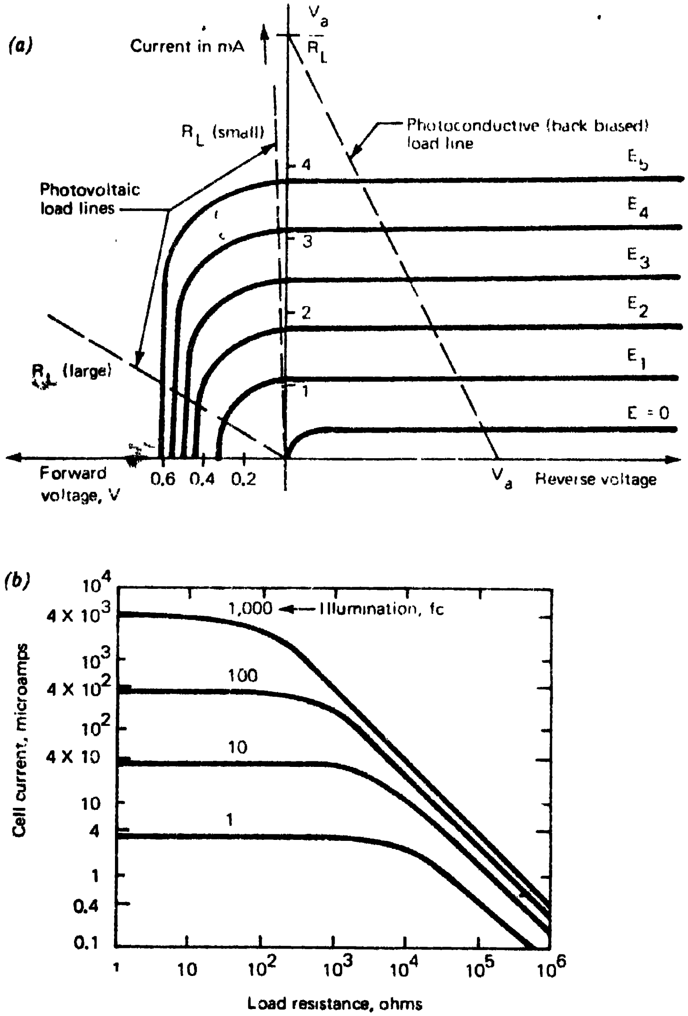


Fig. 27-18. (a) Showing the photoconductive and photovoltaic quadrants of operation for a silicon photodiode. (E = Illumination.) (b) Curves showing the variation in output current from a selenium photovoltaic cell as a function of illumination and load resistance.

This describes a photovoltaic or photogenerative quadrant of operation in which the device can work without any *applied* voltage. That is, the junction *generates* a voltage and/or current depending upon the illumination and the load. An optimum load can be selected that maximizes the output power. (See problem P27-11.)

This mode of operation is very useful when a sensor is required to detect very low levels of radiation, since under completely dark conditions the current is zero. (This is not true in the reverse-biased mode, where the dark current may be comparable to that generated by the radiation being detected.)

Under short-circuit conditions (or R_L very small), the photovoltaic effect is very sensitive to radiation and generates very little noise. Photovoltaic devices employing diffused junctions in indium antimonide and operating at liquid nitrogen temperatures can be used to detect radiation from the visible range to 6 microns, the edge of the far infrared spectrum.

A larger application of the photovoltaic effect is in the use of large-area arrays of flat "sun cells" used to recharge batteries for satellite and earth communication equipment. The open-circuit voltage produced is typically only 0.5 V for either silicon or selenium, and short-circuit cell current is in the order of 1 mA for a 1-square-inch area of Se at 600 footcandles. Thus many such cells must be series and parallel interconnected to provide any significant power. Solar panels providing 5 W at 12 V have been built to permit telemetry systems to operate 24 hours a day by recharging the batteries during daylight hours.

Fig. 27-19 shows a comparison of the selenium and silicon photovoltaic cell construction. The voltage generated is due to the accumulation of carriers produced by photon excitation, as previously described.

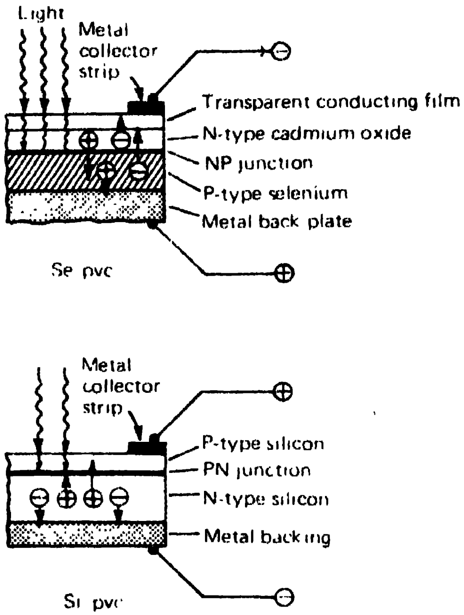


Fig. 27-19. Comparison of Se and Si photovoltaic cell construction.

Because the spectral response of silicon extends well into the infrared (see Fig. 27-20(a)), its efficiency,

$$\eta = \frac{\text{output power}}{\text{input radiation power}},$$

is quite high, approximately 14 percent. Although selenium produces only one-twentieth the output of a silicon cell, it has two advantages. First, it is far superior to silicon in its ability to withstand damaging radiation environments, lasting up to 10,000 times longer than silicon. Secondly, its spectral response approximates very closely that of the human eye (see Fig. 27-20(a)). It is thus used in photometric equipment, exposure meters, and in some cameras to control the automatic exposure diaphragms.

Fig. 27-20(b) indicates a simple selenium footcandle illumination meter in which the current meter presents a very small load. Consequently, the essentially short-circuit

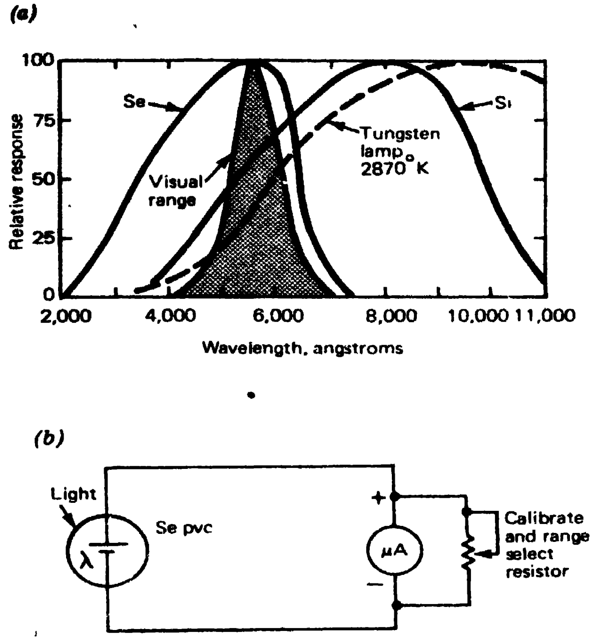


Fig. 27-20. (a) Comparison of Se and Si spectral response curves with the visual range. (b) Simple footcandle meter using a Se photovoltaic cell.

current that flows is directly proportional to the illumination (see the curves in Fig. 27-18(b)), and the meter may be calibrated to read footcandles directly.

27-5. PHOTOEMISSIVE SENSORS

Our discussion of photosensors would not be complete without at least mentioning the photoemissive type. The basic phototubes, vacuum- or gas-filled, rely upon the emission of electrons from a sensitive cathode under the influence of light. The electrons are collected at the anode in a circuit similar to the solid-state photodiode. The cathodes can be made to have different spectral responses, with peaks from 4000 to 8000 Å and sensitivities from 6.5 to 45 μA/lumen. Few applications remain that cannot be filled by their solid-state counterparts.

One photoemissive device that is still very much used is the photomultiplier tube. It can exhibit current amplifications of over 10 million, with luminous anode sensitivities in the hundreds of amps per lumen. This is obtained by means of secondary emission of electrons from *dynodes*, which act as intermediate plates between the cathode and anode. As many as 14 dynodes may be used, each supplied by an increasingly higher voltage to promote secondary emission. Consequently, photomultiplier tubes operate at voltages from 1 to 5 kV. Their ultrahigh sensitivity, low noise, extremely small dark current (as low as 4 picoamps – 4 × 10^{-12}A), and fast response (less than 1 nanosec) make them suitable for use in scintillators and radiation detectors of X-rays, gamma rays, and energetic particles found in nuclear physics. They are also found in space exploration and laser communications.

27-6. LIGHT EMITTERS

Visible LEDs

When current is made to flow in the forward direction through a gallium arsenide phosphide (GaAsP) PN junction, visible light is emitted from the region of the depletion layer. This is due to a recombination of minority carriers so that, when an electron falls from the conduction into the valence energy band, energy is released in a visible form (the mechanism is opposite to that of a photodiode). Such a light-emitting diode is referred to as an LED for short.

Although most LEDs have a peak spectral response at 6600 Å (red), yellow (6100 Å) and green (5600 Å) devices also are available.

High output types used for solid-state lamp indicators (such as the MLED 600) may have a brightness, B_L, of 200 ft lamberts (a measure of source intensity) at only 10 mA and 1.6 V. This represents only 16 mW of input power. The brightness may be increased to 1100 ft lamberts at 50 mA. LEDs used as indicators have solid-state reliability of typically 10^6 hours lifetime. They may also be used in optical switching circuits, since, unlike an incandescent lamp, they can be switched on or off in 1 nsec. Before looking at other applications, let us briefly refer to two other solid-state emitters.

Infrared Emitters

If the PN junction results from a diffusion in gallium arsenide (GaAs), the radiant energy is infrared with a typical peak at 9000 Å. This ideally matches the response of silicon photodiodes and phototransistors. For example, in the shaft encoder of Fig. 27-15, the MLED 910 can produce 150 μW of radiant energy at 50 mA and 1.2 V. Its diameter is 0.06 inch and length 0.125 inch, so it may be used in high-density mounting applications.

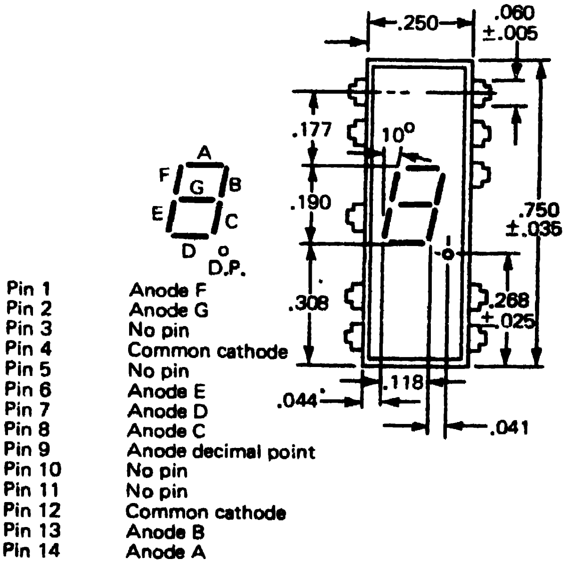
Laser Diodes

The energy emitted from visible LEDs and IR emitters is spontaneous or noncoherent (random in phase and direction). If the level of current flowing through a GaAs PN junction is increased beyond a certain threshold, a *lasing* action takes place, in which the spontaneous emission *stimulates* an increase in the radiant power. This *l*ight *a*mplification by the *s*timulated *e*mission of *r*adiation (LASER) is something similar to an avalanche effect. The resultant output from a solid-state laser is invisible and is not as coherent as that produced from a ruby, gas, or dye laser, but the device is much more compact. Output radiant powers from 1 to 65 W are possible from single-diode lasers when pulsed for 0.2 μsec with current from 10 to 250 A.

27-7. ALPHANUMERIC DISPLAYS

One way of producing an alphanumeric display is to make a 5 × 7 array of LEDs (plus one for the decimal point) in a monolithic structure. By driving the appropriate horizontal and vertical inputs, 64 different characters can be displayed, including the full alphabet, all numerals, and various signs.

Another method that can display all numerals and nine letters is the seven-segment monolithic device shown in Fig. 27-21. Mounted in a standard DIP package, each of the seven segments consists of five LEDs, interconnected by evaporated aluminum on the surface of the chip. Each segment provides a brightness of 200 ft lamberts at 5 mA and 1.7 V. The compactness of the display allows numerous packages to be mounted side



Pin 1	Anode F
Pin 2	Anode G
Pin 3	No pin
Pin 4	Common cathode
Pin 5	No pin
Pin 6	Anode E
Pin 7	Anode D
Pin 8	Anode C
Pin 9	Anode decimal point
Pin 10	No pin
Pin 11	No pin
Pin 12	Common cathode
Pin 13	Anode B
Pin 14	Anode A

Fig. 27-21. Showing a 0.19 inch high LED, seven-segment display mounted in a standard DIP package. (Courtesy of Monsanto Co.)

by side to give any degree of readout. In the electronic calculator in Fig. 24-22, five such displays, each 0.112 inch high by 0.062 inch wide, in one 14 pin DIP package provides even greater mounting density.

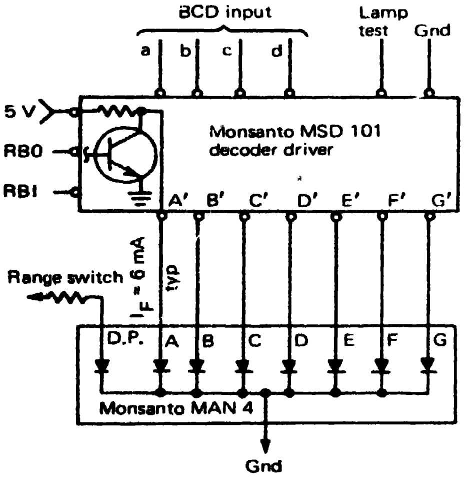
As mentioned in Chapter 26, Fig. 26-15, the seven-segment displays are driven from a decoder/driver that accepts a BCD input. A functional diagram of a Monsanto MSD 101 decoder/driver connected to a MAN 4 display unit is shown in Fig. 27-22, along with the truth table. Some suppliers can provide this whole function on one monolithic LSI chip.

Another type of display that uses a seven-segment system is the *electroluminescent* or *liquid crystal* display. This consists of a thin layer of normally transparent liquid crystal material between two electrodes. When an electric field is applied (15 to 60 V), the material becomes turbulent, reflecting and scattering ambient light. It provides excellent brightness under high ambient light conditions and requires only 50 μW of power per segment. This power is much lower than for the LED, but the life expectancy is not as high as the LED (minimum 10,000 hours).

27-8. PHOTOCOUPLERS

The fact that the light output from an LED is proportional to the input current makes it an ideal device for coupling one circuit to another by means of a photosensor. When the emitter and sensor are placed in the same light-tight package they are known as a coupled pair or optoisolator. The sensor may be a PIN photodiode for fast switching, which can work in the photoconductive or photovoltaic mode, as shown in Fig. 27-23. Or it may be a phototransistor or photo SCR. Typical isolation resistance between emitter and detector is 10^{11} ohms with isolation voltages up

(a)



(b)

LT	RBI		BI/RBO	Display
0	X		1	8
X	X		0	Blank
1	0		0	Blank
1	1		1	0
1	X		1	1
1	X		1	2
1	X		1	3
1	X			4
1	X			5
1	X			
1	X			
1	X			8
1	X			9
1	X			-
1	X			Blank
1	X			H
1	X			L
1	X			P
	X	1		Blank

X = Don't care condition

Fig. 27-22. (a) Functional block diagram of the decoder/driver connected to the seven-segment LED display. (b) Truth table of the decoder driver showing the output display. (Courtesy of Monsanto Co.)

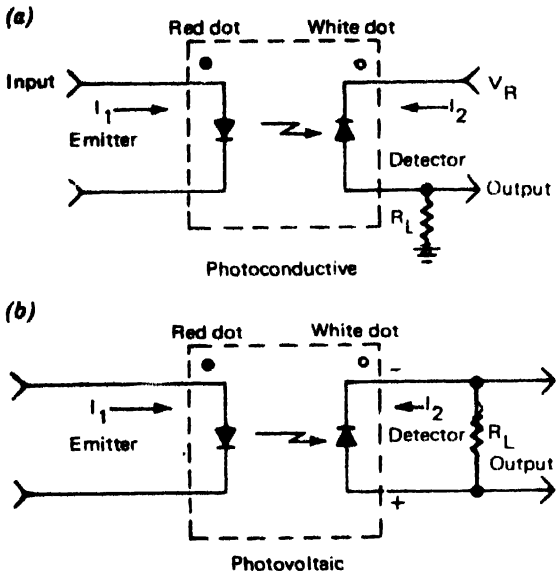


Fig. 27-23. Photodiode-coupled pair operating in the (a) photoconductive or (b) photovoltaic modes.

to 2500 V. Output current in phototransistor-coupled pairs is typically 35% of the LED input current with a linear relation between the two. GaAs LEDs are used to provide spectral matching with the silicon sensors.

In some applications, the light may be directed from a single remote source onto two or more photosensors simultaneously by means of *fiber optics*. These are flexible, often transparent, "light pipes" that transmit energy (from 4000 Å to the infrared range) by means of total internal reflection. Single fiber optics may be as small as 0.0017 inch in diameter. Even smaller diameters may be combined in coherent *bundles* for image transmission applications. Approximately 50% of the input light can be transmitted over lengths up to 3 ft long. Various combinations of fiber optics are used in applications ranging from auto dash lighting to microfilm reading, from numerical displays to miniature oral dental lighting units.

The field of optoelectronics is growing rapidly, with new applications being found every day.

27-9. SUMMARY

(1) Visible light occupies only a small portion of the electromagnetic spectrum; the human eye responds to radiation that extends from 0.4 to 0.76 microns in wavelength.

(2) The radiation from a tungsten lamp is mostly in the infrared region, with relatively little in the visible region. It is often used as a source of infrared radiation.

(3) Illumination is a measure of the *visible* radiation on a surface and is measured in footcandles or lumens/sq ft.

(4) Irradiance is a measure of the *total* radiation on a surface and is measured in watts/cm^2.

(5) An illuminance (illumination) of 1 lumen/sq ft occurs on a surface placed one foot away from a source of 1 candlepower or 4π lumens.

(6) Bulk-type photoconductive cells are constructed from cadmium sulfide or cadmium selenide semiconductors. They have their conductivity increased by incident photons of sufficient energy raising valence electrons to the conduction band.

(7) Bulk-type photocells are characterized by large power dissipation, slow switching speed and peak response at various wavelengths in the visible spectrum.

(8) Junction-type photoconductors are photodiodes or phototransistors wherein one or more PN junctions are used under reverse bias. Electron-hole pairs resulting from photons incident on the PN junction add to the minority carriers due to the reverse bias

(9) Silicon and germanium photosensors have their peak spectral response in the infrared region, with their response in the visible only approximately 40% of maximum.

(10) Variations of the basic PN photodiode are: (a) PIN photodiode (ultrafast response), (b) avalanche photodiode (higher current sensitivity and fast response), (c) photoduodiode (inexpensive higher current sensitivity but slower response).

(11) The inherent current gain of a transistor provides the phototransistor with a high current gain in the region of 1 mA/mW/cm^2. By adding a second conventional transistor, a photo Darlington amplifier results, with even higher sensitivity.

(12) Photodiodes and phototransistors may be operated as switching devices in light-operated relays, shaft encoders, paper tape readers, brushless dc motors, etc., or used to provide an output proportional to the light input as in the accelerometer.

(13) Photovoltaic sensors, usually either silicon or selenium, generate a voltage under open circuit conditions of typically 0.5 V. When operated in the short-circuit mode, the current is proportional to the illumination and can be used to provide a direct-reading foot-candle meter.

(14) Photovoltaic sensors may be series-parallel interconnected to provide large-area solar power cells.

(15) Photoemissive sensors are vacuum or gas-filled phototubes. The only device having real advantages over the solid-state sensors is the photomultiplier tube, which has current amplifications of over 10 million and extremely low-noise operation.

(16) Light emitting diodes, LEDs, may emit visible radiation when the PN junction is diffused in GaAsP, infrared if constructed from GaAs, and a laser beam if sufficiently highly pulsed current is applied to a GaAs junction.

(17) A popular alphanumeric display is the LED seven-segment device, which can provide high brightness for only a few milliwatts of input power per segment. A liquid crystal display requires only a few microwatts but has a lower life expectancy.

(18) A photocoupled pair, such as an LED and a phototransistor in the same package, provides an output current proportional to input current, with a high degree of isolation. Fiber optics (light pipes) may also be used to couple source and sensor over relatively large distances.

QUESTIONS

Q27-1. (a) What is the electromagnetic spectrum?

(b) What four units are used as a measure of wavelength?

(c) State the range of wavelengths in the visible region in the four units in (b).

(d) What is a photon?

(e) Do X-rays and gamma rays have a higher or lower wavelength than visible radiation?

(f) Explain why these rays should be handled with care, with reference to Eq. 27-2.

Q27-2. (a) Explain, with reference to its spectral response curve, why a tungsten incandescent lamp is a relatively inefficient device to provide illumination.

(b) What happens to the output radiation from a tungsten lamp as its operating temperature is reduced?

Q27-3. (a) What units are used to measure illuminance (illumination) and radiance? Distinguish carefully between the two terms.

(b) Is it possible for a source of radiation to provide irradiance but no illuminance?
(c) Is it possible for a source to provide illuminance but no radiation?
(d) Which is the more efficient illuminator, a 100-watt tungsten incandescent lamp or a 100-watt fluorescent lamp? Why?

Q27-4. (a) Describe the mechanism by which light increases the conductivity of a bulk-type photoconductive cell.
(b) What are the usual materials found in these cells?
(c) Give at least three characteristics of bulk-type photo cells. Which is considered a drawback?

Q27-5. (a) Describe the mechanism by which light (radiant energy) controls the flow of current in a reverse-biased photo junction.
(b) Would the junction current be responsive to light if forward biased?
(c) How does a PIN photodiode differ from a conventional PN photodiode? What advantageous characteristic results from this modification?
(d) How does an avalanche photodiode obtain a higher current sensitivity than a PIN photodiode? Does it have fast response? What is its drawback?

Q27-6. (a) What does the name photoduodiode imply?
(b) How does an NPN photoduodiode differ from a phototransistor?
(c) What is the main advantage of a phototransistor?
(d) What solid-state device has an even higher sensitivity than a phototransistor? What is its main drawback?
(e) Where is the peak spectral response of silicon photojunction devices?
(f) May a tungsten lamp be used as the light source? What disadvantages might it have?

Q27-7. If a phototransistor were used in place of the photoconductive cell in Fig. 27-7(c), and you had the collector characteristics of the phototransistor, describe how you would draw the load line on the curves and determine the irradiance necessary to ensure that the output transistor would be driven into saturation to activate the relay (R_1 and V_{CC} given).

Q27-8. (a) In the code wheel for the electro-optical shaft encoder in Fig. 27-15, indicate the position number of each pie-shaped sector. (Note the binary code for position 5.)
(b) Does the outermost or innermost "ring" in the code wheel constitute the MSB (most significant bit)?
(c) Explain how the presence of light (a hole) provides a "0" output, but no light gives a "1."

Q27-9. In the brushless motor of Fig. 27-16, the dotted lines representing the permanent magnet rotor correspond to the dotted lines in the light shield after winding A has been energized. Describe the sequence of operations for a complete revolution when power is first applied with the rotor in the position shown solid.

Q27-10. The phototransistor of the accelerometer in Fig. 27-17(a) is being used in a linear circuit to produce an output voltage proportional to the light input.

(a) For the circuit shown, what must be the at-rest position of the shutter with respect to the light beam with no acceleration to obtain the curve in (c)?
(b) Why does the curve show a flattening effect at high g's?
(c) Describe, by drawing a sketch, how the output voltage would vary if the device were placed in an elevator on the ground

floor, then accelerated constantly at 2 g's for 4 floors, run at constant speed for 6 floors, then decelerated constantly at 2 g's for 4 floors to come to rest.

Q27-11. (a) What is meant by the photovoltaic mode of operation for a photodiode?
(b) When is it possible that this type of operation would be useful as a sensor?
(c) Which cell is more efficient in converting radiant energy to electrical power?
(d) What advantage does selenium provide?
(e) Why is selenium popular for use in exposure meters and illumination meters?

Q27-12. (a) What advantages does a photomultiplier tube offer over a solid-state sensor?
(b) To what family of photosensors does this device belong?
(c) What do you understand by the word "dynode"?

Q27-13. (a) If LEDs are available that emit in red, yellow, and green colors, what can you conclude regarding their spectral response?
(b) What is the brightness of a visible LED measured in?
(c) What is the output of an infrared diode measured in?
(d) What are some advantages of an LED as a lamp indicator?
(e) What does each segment of the seven-segment alphanumeric display consist of?
(f) Which segments would you activate to form the letters E and F? What other distinct letters could be formed in the MAN 4 display?

Q27-14. (a) What is a liquid crystal display?
(b) What is its advantage compared with the LED?
(c) What is at least one disadvantage?

Q27-15. (a) What combinations of photocoupled pairs are possible?
(b) When in general is it advisable to use a coupled pair? Give two practical examples.
(c) Give two of your own examples where fiber optics might be useful.

PROBLEMS

P27-1. (a) Determine the range of frequencies occupied by the visible spectrum.
(b) Determine the energy of photons at each extreme of the visible spectrum.
(c) Determine the energy of high-frequency X-rays at a wavelength of 0.5×10^{-3} millimicrons and compare with the energy of light in (b).

P27-2. (a) Use the curves in Fig. 27-3 to determine the relative response of the eye (expressed as a percentage of maximum) to red, amber (yellow), and green.
(b) What do these figures suggest concerning the relative ease with which traffic lights can be seen, if each is powered by the same size of tungsten lamp?

P27-3. A 40 W point source of light emits 160π lumens of light flux.

(a) Assuming the light radiates outward equally in all directions, determine the illuminance on a surface in footcandles at distances of 1 ft, 2 ft, 3 ft, 4 ft, and 5 ft from the source.
(b) Sketch a curve showing how the illuminance varies with distance. What is the name of this kind of curve?

P27-4. (a) Assume that a reflector is used on a 100-watt tungsten lamp (1630 lumens) to direct the radiation in a hemisphere. Determine the illuminance on a surface in footcandles at integral distances of 1 ft from 1 ft to 5 ft from the source.
(b) Repeat (a) for a 100-watt fluorescent lamp (4,400 lumens).

(c) Sketch curves of illuminance vs. distance from (a) and (b) and comment.

P27-5. The irradiance, H, upon a surface located at a distance d cm from an LED of output power P watts and having a spatial beam angle of θ radians is given by

$$H = \frac{4P}{\pi\theta^2 d^2} \text{ watts/cm}^2.$$

Determine the irradiance at distances of 1, 2, 3, 4, and 5 cm from an MV50 economy visible LED that provides an output radiated power of 38 μW at a beam angle of 60°. (180° = π radians.)

P27-6. Repeat P27-5 for an ME5 high-power infrared emitter that radiates 18 mW at a beam angle of 30°.

P27-7. A photoconductive cell that has the characteristics given by curve E in Fig. 27-5 is used in the circuit of Fig. 27-7(a). The relay has a resistance of 100 ohms and operates at a current of 50 mA. The voltage source is 12 V dc. If the cell is placed outside, determine approximately:

(a) How much current flows in the circuit at full moon (0.01 footcandle).
(b) How much current flows in the circuit at twilight (1 footcandle).
(c) The cell illumination necessary to operate the relay.
(d) The power dissipation of the cell for the above three illuminations.

P27-8. A 120 V ac/dc universal motor that runs at full speed of 6000 rpm with a current of 0.22 A is used in the circuit of Fig. 27-7(b). The photoconductive cell obeys curve E in Fig. 27-5.

(a) Assuming the speed of the motor is directly proportional to the current through the motor, determine its top speed if the maximum illumination provided by the lamp on the cell is 100 footcandles.
(b) What illumination will reduce the motor speed to approximately one-third of full rated speed? (Hint: Determine the effective resistance of the motor and assume it stays constant.)
(c) Determine the approximate dissipation of the cell for (a) and (b). What will it be necessary to do to the cell for this type of operation?

P27-9. The FPT/120 phototransistor (Fig. 27-13) is being used as an optical logic inverter to provide a 0 (or low voltage) when light is present and a 1 (higher voltage) when light is not present. The output is taken at the collector as in a conventional inverter. If V_{CC} = 12 V and R_L = 1.2 K ohms, determine the output voltages for 0 and 1 if the illumination is 100 footcandles and 0 footcandles, respectively, from a tungsten source at 2870°K. (Hint: Draw the load line on the curves, one end at V_{CC}, the other at V_{CC}/R_L.)

P27-10. In order to improve the optical inverter in P27-9 the same transistor is operated with the same voltage but R_L = 5.6 K ohms.

(a) What will be the output voltage at 100 footcandles?
(b) What will be the output voltage at 0 footcandles?
(c) Given that the phototransistor's sensitivity is 2.5mA/mW/cm^2, determine the maximum irradiance and illumination that can be permitted if a 1 output is to correspond to 11.9 V.
(d) If the infrared emitter in P27-6 is to provide the radiation for the output in (a), what must be the maximum separation between emitter and sensor?

(e) Repeat (d) for the emitter in P27-5. (In practice, a correction factor would be necessary because of the different spectral responses involved, requiring the separation to be even smaller.)

P27-11. (a) On the photodiode curves of Fig. 27-18(a), draw photovoltaic load lines for: (1) $R_L = 50$ ohms, (2) $R_L = 100$ ohms, (3) $R_L = 1$ K ohm, and determine the load power in R_L at maximum illumination for each case.

(b) Sketch a curve showing how output load power varies with load resistance and determine the optimum value of R_L for maximum output power. (Draw more load lines if necessary.)

28

THYRISTORS AND THE UNIJUNCTION TRANSISTOR

28-0. INTRODUCTION

Thyristors comprise those solid state devices that have two or more PN junctions (three, four, or five semiconductor layers) which can be switched from a nonconducting to a conducting state. The name is derived from a ***thyratron*** which may be thought of as the tube counterpart of the silicon controlled rectifier (SCR) although this is only one member of the thyristor family.

Fig. 28-1 shows that there are two main divisions in the thyristor family. The SCR is a unidirectional device allowing current to flow in only one forward direction, while a triac is of the bidirectional type and may be thought of as two SCRs connected in anti-parallel.

Once a thyristor "fires," current continues to flow until the voltage across the device reverses. In order to trigger the thyristor into conduction, a number of triggering devices have been developed, which, with suitable circuitry, can control the point at which the thyristor begins to conduct. In this way, ac and dc power variation is made possible allowing light, heat, and motor speed control to be carried out economically and efficiently

In this chapter, we examine the characteristics of power thyristors and their related triggering devices. Also included is the unijunction transistor, since it is ideally suited for firing thyristors when connected in a suitable relaxation oscillator circuit.

28-1. THE SILICON CONTROLLED RECTIFIER (SCR)

Construction

As indicated in Fig. 28-1, an SCR is a four-layer PNPN device. One type of construction consists of alloying an N-type material into a silicon PNP pellet, as shown in Fig. 28-2(a). This provides the main two terminals, the anode and cathode. The gate contact is then welded to the P region. Undoubtedly, it was this type of construction that led to the symbol shown in Fig. 28-2(b). However, there is

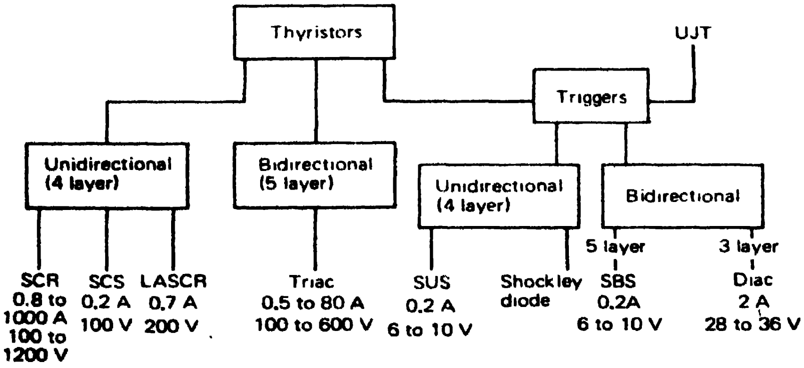

Thyristors
UJT
Triggers
Unidirectional (4 layer)
Bidirectional (5 layer)
Unidirectional (4 layer)
Bidirectional
5 layer
3 layer
SCR 0.8 to 1000 A 100 to 1200 V
SCS 0.2 A 100 V
LASCR 0.7 A 200 V
Triac 0.5 to 80 A 100 to 600 V
SUS 0.2 A 6 to 10 V
Shockley diode
SBS 0.2A 6 to 10 V
Diac 2 A 28 to 36 V

Fig. 28-1. Diagram showing the thyristor family.

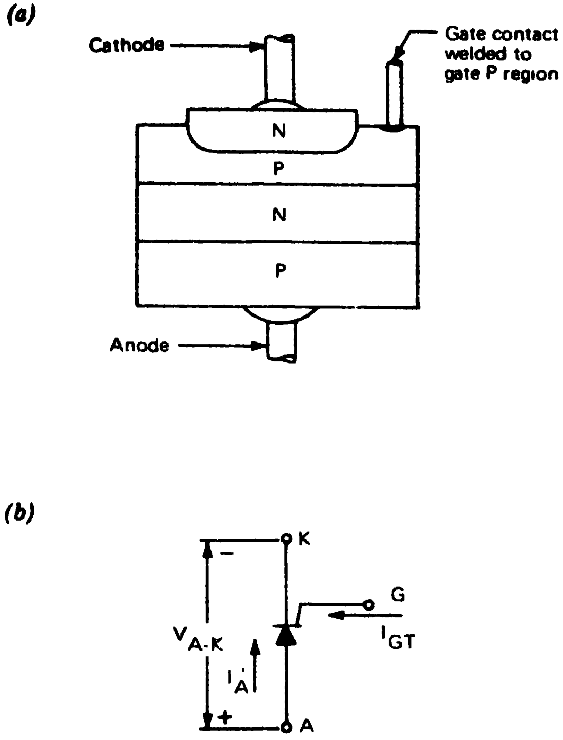


Fig. 28-2. (a) One type of construction used in an SCR that clearly shows the four PNPN layers. (b) Standard symbol of an SCR.

an annular type of construction in which the gate is central to the device, with the cathode electrode surrounding it. This is called the shorted-emitter construction, because of the partial overlay of the cathode electrode onto the P-type gate. Fast turn-on and improved high-temperature capabilities are among the claims made for this type of construction.

If a planar type of construction is employed and connection is also made to the other N-type region, a four-terminal device called a *silicon controlled switch* (SCS) results.

An LASCR is a *light-activated SCR* and consists of mounting the PNPN pellet in a package provided with a translucent window. Although connection to the gate may also be provided, light energy or radiant energy on the PN junction is responsible for turning the device on.

Static Characteristics

The characteristic curves in Fig. 28-3 show how the gate current into the SCR determines the necessary voltage from anode to cathode to cause conduction.

If the gate is open ($I_G = 0$), with forward voltage applied from anode to gate, the SCR is in the forward blocking region or off state. If, however. the voltage exceeds the forward breakover voltage V_{BO} (abbreviated from $V_{F(BO)O}$), the SCR will turn on, and current will be limited only by the applied voltage and the resistance in series with the device. At the breakover, the characteristic curve is shown dotted, since this is an unstable, fast-switching region. A parameter often used to describe the repetitive peak forward off-state voltage that an SCR can withstand (with the gate open) is V_{DROM}. This may be equal to or larger than V_{RROM}, (PRV), the repetitive peak reverse voltage (open gate) that the SCR can be subjected to before reverse breakdown.

As gate current is increased, the forward breakover voltage decreases, until for a very large value of I_G (in the order of 50 mA) the SCR turns on immediately when voltage is applied.

The normal mode of operation is for the voltage to exist across the SCR and then for the SCR to be turned on by a pulse of current in the gate. The voltage drop across the diode is typically 1.5 V at its rated current. Once the diode is on, a minimum amount of holding current, I_H, is required to flow to keep the device on. If the current is lowered below I_H by increasing the external circuit resistance, the SCR will switch off.

It should be noted that once the SCR is fired into conduction the gate loses control.

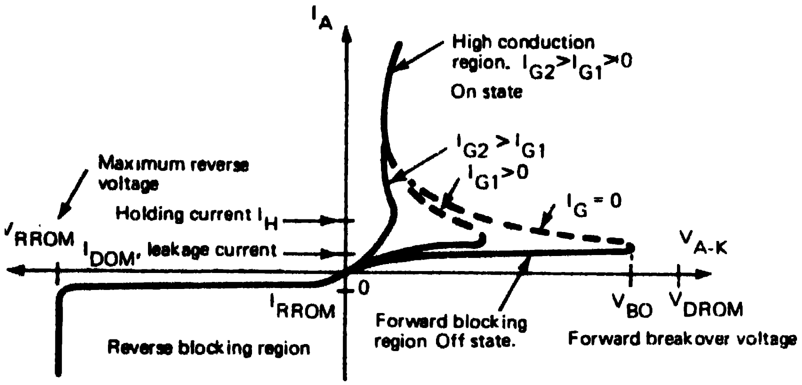


Fig. 28-3. Characteristic curves of an SCR showing how the forward breakover voltage deperds upon the value of gate current.

That is, the gate cannot be used to switch the device off. Only if $V_{A\text{-}K}$ is reduced to zero (as automatically happens in rectifying applications), or if the current is reduced below I_H will the device turn off. (The SCS, a much lower power device (0.2 A at 100 V PRV) *is* capable of being switched off, with its second gate connection aiding in this function.)

Analysis of the Operation of the SCR

The SCR may be analyzed by using a two-transistor analogy in which the SCR is "split down the middle," as in Fig. 28-4.

Consider a positive voltage on the anode with respect to the cathode, no gate signal applied. The anode current is then just the sum of the leakage currents of the two "transistors," $I_{CO1} + I_{CO2}$.

Remember the basic current equation for a CE transistor:

$$I_C = \beta I_B + (\beta + 1) I_{CO} \tag{28-1}$$

$$= \beta I_B + \beta I_{CO} + I_{CO} \tag{28-2}$$

$$= \beta(I_B + I_{CO}) + I_{CO} = h_{FE}(I_B + I_{CO}) + I_{CO}. \tag{28-3}$$

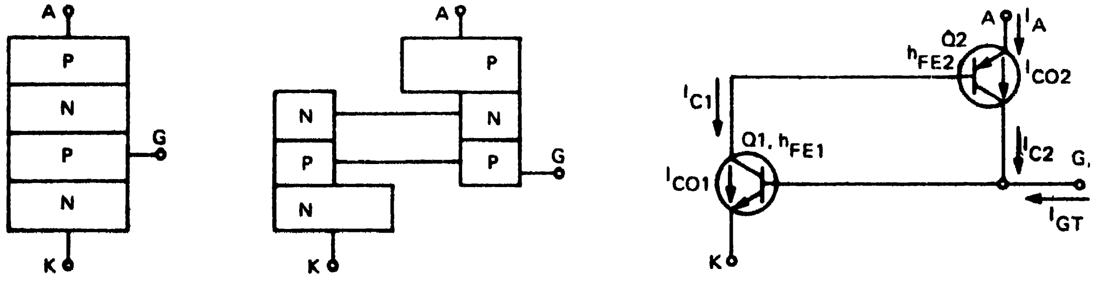


Fig. 28-4. Showing how the SCR may be thought of as two interconnected transistors with regenerative feedback.

Consider Q1: Note that the base current, I_B for Q1 is the collector current of Q2, I_{C2}.

$$I_{C1} = h_{FE1}(I_{C2} + I_{CO1}) + I_{CO1}. \qquad (28\text{-}4)$$

Consider Q2: Note that the base current I_B for Q2 is the collector current of Q1, I_{C1}.

$$I_{C2} = h_{FE2}(I_{C1} + I_{CO2}) + I_{CO2}. \qquad (28\text{-}5)$$

Then $I_A = I_{C1} + I_{C2}$

$$= \frac{(1 + h_{FE1})(1 + h_{FE2})(I_{CO1} + I_{CO2})}{1 - h_{FE1} \times h_{FE2}} \qquad (28\text{-}6)$$

With no gate current applied, let us verify that this equation will give us our earlier result that $I_A = I_{CO1} + I_{CO2}$.

Due to the intentional "wide base regions" of each "transistor," the h_{FE1} and h_{FE2} are very low (much less than one) when no gate current is applied. (That is, when $I_E \approx 0$, h_{FE} is near zero for a silicon transistor. See Fig. 9-21.) Thus, the loop again, $G = h_{FE1} \times h_{FE2}$, is much less than unity, and in the equation for I_A we have:

$$= \frac{(I_{CO1} + I_{CO2})(1 + 0^+)(1 + 0^+)}{1 - 0^+ \times 0^+}$$

$$\approx I_{CO1} + I_{CO2}. \qquad (28\text{-}7)$$

0^+ represents a quantity slightly greater than 0. Thus the SCR is off and provides high blocking.

Now consider applying a positive voltage from gate to cathode to inject holes into the base of Q1. This forward biases that base-emitter junction, increasing I_{C1}. But this collector current is the base current for Q2, so "transistor" Q2 is turned on, increasing its collector and emitter current. The collector current of Q2 can now take the place of the trigger current and turn Q1 on more. The action is one of internal regenerative feedback, until both "transistors" are driven into saturation, all junctions are forward biased, and the anode current I_A is limited only by the external circuit resistance and voltage. Removal of I_{GT} after I_A has reached the latching current (equal to $3 \times I_H$) does not stop conduction, so just a pulse of current is required to trigger the SCR on.

Now how does the equation explain this? It was shown earlier, in Fig. 9-21, that the h_{FE} of a silicon transistor is very low when the collector current is small but increases rapidly when I_C increases. In our "transistors, h_{FE1} and h_{FE2} approach 1 (unity) when I_C increases due to the gate current pulse and the succeeding buildup of current. Thus, in the equation for I_A:

$$I_A = \frac{(1 + 1^-)(1 + 1^-)(I_{CO1} + I_{CO2})}{1 - 1^- \times 1^-}$$

$$\frac{4(I_{CO1} + I_{CO2})}{0^+} \approx \infty. \qquad (28\text{-}8)$$

1^- represents a quantity slightly less than 1.

Of course, the current cannot be infinite in practice but increases to a limit determined by circuit conditions.

The SCR can also be triggered on in the following ways:

(1) High voltage from anode to cathode (greater than V_{BO}) can increase the energy of the leakage current carriers, I_{CO}, and cause an avalanche effect, causing

$$h_{FE1} \times h_{FE2} \to 1.$$

(2) Due to the capacitance of PN junctions, a rapid rate of rise of voltage from anode to cathode, dv/dt, will cause a charging current $i = C\,dv/dt$, which can make $G \to 1$.

(Typical critical rate of rise of voltage at $T_C = 100°\text{C}$ is 200 V/μsec.)

(3) Temperature rise will make the leakage current increase, so that G → 1.

(4) Radiant energy, such as light, can trigger the LASCR. The first three "methods" are not used to purposely trigger the SCR. A good SCR is designed to avoid being triggered on in these ways, or circuit design is modified to minimize these problems.

Gate Triggering Characteristics

The dc gate trigger current, I_{GT}, is the current necessary to trigger the SCR on when the anode is typically 6 to 12 V dc with respect to the cathode; V_{GT} is the gate voltage with respect to the cathode just before conduction starts.

Since triggering requirements for SCRs of the same type vary from unit to unit, the manufacturer provides forward gate characteristics including the range of values of V_{GT} and I_{GT} that one might find. As can be seen from the curves in Fig. 28-5, the triggering characteristics vary with temperature. It requires a higher value of gate current and voltage to fire an SCR at a lower temperature. For example, for the types shown in Fig. 28-5, a maximum V_{GT} of 2 V and a maximum I_{GT} of 40 mA will ensure that all units will fire at 25°C junction temperature. However, at -40°C, it will require a maximum V_{GT} of 3 V and a maximum I_{GT} of 80 mA.

The values given for I_{GT} and V_{GT} in this example are indicated as the *maximum* values

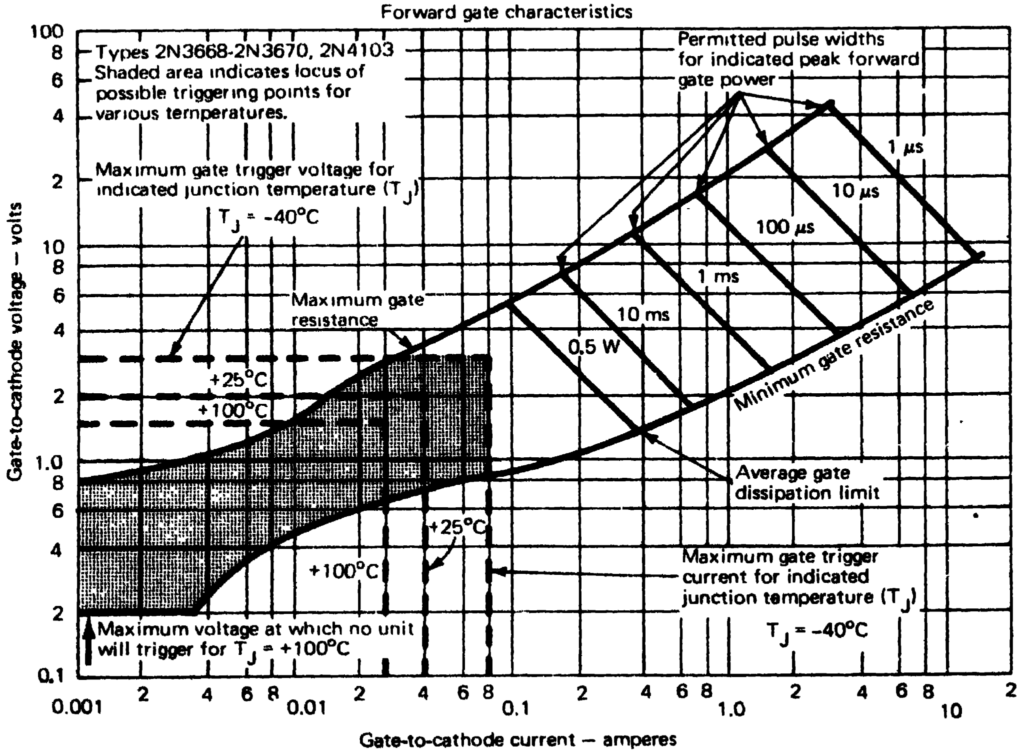


Fig. 28-5. Forward gate characteristics of 2N3668-2N3670 SCRs showing the preferred areas of gate triggering. (Courtesy of RCA Corp.)

needed. It is quite possible that triggering will occur at lower values. For example, at 25°C, typical values for V_{GT} and I_{GT} are 1.5 V and 20 mA respectively. Consequently, to ensure reliable switching for all units, the preferred region for triggering lies between the two curves of minimum and maximum gate resistance, to the right of the dotted lines in the shaded region (corresponding to the anticipated temperature) and to the left of the maximum dissipation curves.

The gate dissipation of the 2N3668-2N3670 SCRs is listed as $P_{G(AV)} = 0.5$ W, $P_{GM} = 40$ W for 10 μsec. That is, if the gate is triggered from a dc source, a continuous current of only 0.2 A at $V_{GT} = 2.5$ V, for example, could be safely dissipated. If, however, the gate current is supplied in a narrow pulse, of 10 μsec duration for example, the gate current could be made as high as 2 A at a gate voltage of 20 V without exceeding the peak transient dissipation, P_{GM}, of 40 W.

What is the purpose of using such a high gate current pulse if the SCR will switch on for a much lower trigger power? The answer is that the SCR will switch on much more quickly with a higher gate current. For example, the gate turn-on time, t_{gt}, is typically 1.25 μsec if $I_{GT} = 0.2$ A, increasing to 5 μsec if I_{GT} is only 0.02 A.

Since the minimum duration to trigger an SCR is in the order of 1 to 10 μsec, it is not unusual to find high gate currents being used, although this requires a very low internal resistance in the gate triggering voltage source. (See P28-4.) The fast turn-on of the SCR is not without its problems. If the rate of rise of the on-state current exceeds the rated critical value of di/dt (typically 200 A/μsec), damaging localized hot spots may occur in the device.

Simple Applications of the SCR

Before we can look at some of the typical current ratings of SCRs, we must look at some basic applications of the device to see the current waveforms involved.

Variable Half-Wave Rectifier. Fig. 28-6(a) shows a triggering circuit that consists simply of a diode and resistor in series with the gate. We will consider more suitable triggering circuits later wherein D1 is replaced by a three- or four-layer diode. In this application, D1 serves to block the reverse voltage on the gate when the applied ac voltage reverses.

By increasing R, the point on the sine wave at which the SCR fires can be made to occur later, up to a maximum of 90°. The SCR automatically stops conducting when the voltage reverses. Clearly, the limits on the

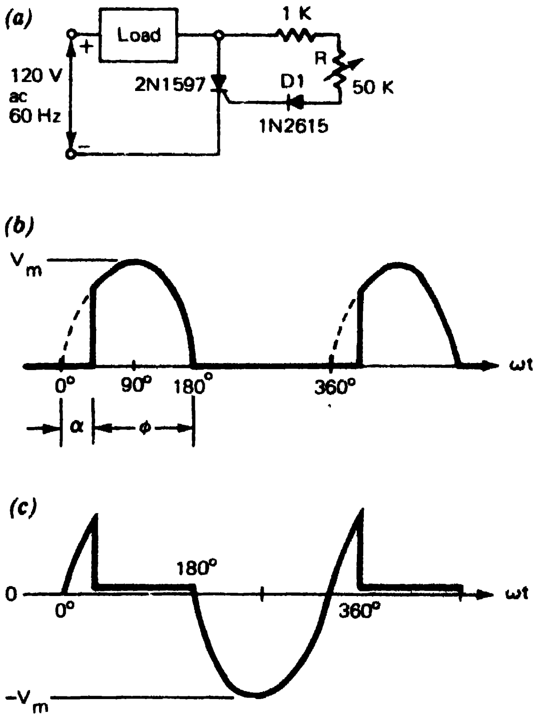


Fig. 28-6. (a) Circuit of a variable half-wave SCR rectifier. (b) Waveform of load voltage showing the firing angle α and the conduction angle φ. (c) Waveform of the SCR voltage, $V_{A\text{-}K}$.

firing and conduction angles in Fig. 28-6(b) are:

Firing angle α: $0° \leq \alpha \leq 90°$. (28-9)

Conduction angle ϕ: $90° \leq \phi \leq 180°$. (28-10)

The blocking action of the SCR in the reverse direction and partially in the forward direction is shown in Fig. 28-6(c).

180° Variable Half-Wave Rectifier. The addition of a phase-shift circuit, as in Fig. 28-7(a) effectively delays the point at which the SCR will trigger on the positive half of the sine wave. The charging of C through R provides a certain delay, so that the firing angle can be controlled up to a maximum of 180°. In this case the circuit is always off, and no current flows at all. The ranges of firing and conduction angles for this circuit are now given by:

Firing angle α: $0° \leq \alpha \leq 180°$. (28-11)

Conduction angle ϕ: $0° \leq \phi \leq 180°$. (28-12)

Fig. 28-7(b) shows that this control is made possible by the phase shift introduced by the capacitor to make v_C lag behind v_i. That is,

$$\frac{V_i}{1 + j\omega CR} \qquad (28\text{-}13)$$

If $\omega CR \ll 1$, V_C and V_i are practically in phase, and firing occurs almost immediately after v_i is positive ($\alpha = 0$, $\phi = 180°$). If $\omega CR \gg 1$, V_C is delayed by practically 90° from V_i and is much reduced in amplitude. Thus, firing may be delayed to almost the end of half a cycle ($\alpha = 180°$, $\phi = 0°$).

Diode D2 is sometimes added to reset C by discharging it during the negative half cycle, so that firing will occur at the same point each cycle.

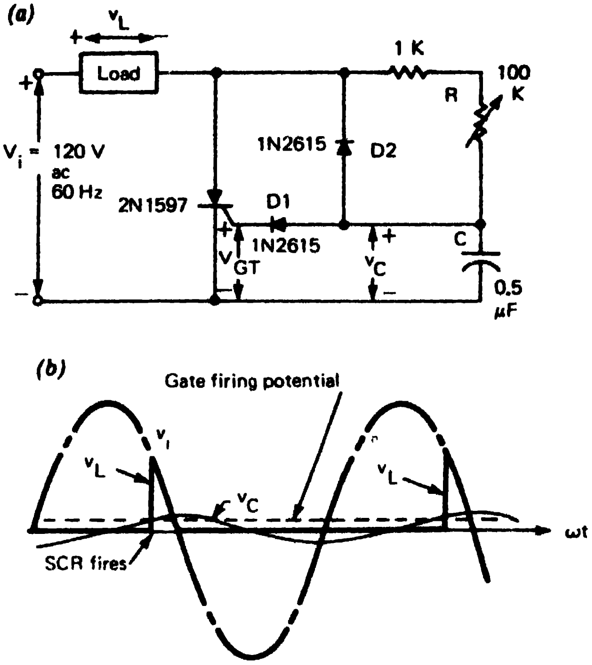


Fig. 28-7. (a) Circuit of a 180° variable half-wave rectifier. (b) Voltage waveforms in the rectifier circuit showing v_C lagging v_i.

Variable Full-Wave Rectifier. One method of obtaining full-wave rectification control is shown in Fig 28-8(a). The gate control input has been left open, since a number of different firing circuits could be applied. As shown by Fig. 28-8(b), the four diodes provide full-wave rectification, with the SCR controlling the amount applied to the load.

All of the previous circuits have been able to control the average value of current through the load. Thus, they can be used to vary the brightness of a lamp, the speed of a dc motor, or the heat developed in an electrical heater, etc. It is to be noted that this is accomplished without using a large power rheostat, so it is a very efficient method.

Current Ratings of the SCR

For the half-wave rectifier circuits in Fig. 28-6 and 7, the average current flowing through

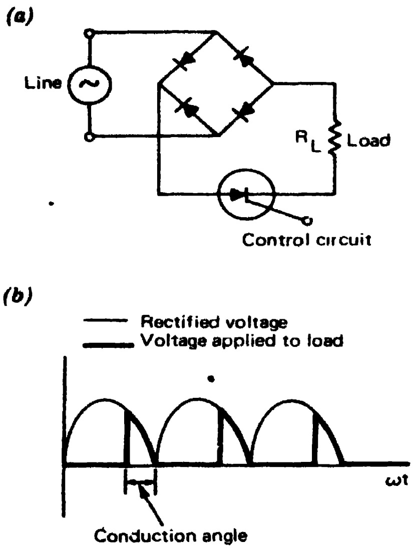


Fig. 28-8. (a) Circuit to provide SCR control of a full-wave rectifier. (b) The rectified and load voltage waveforms.

the device may be expressed in terms of V_m, R_L, and the firing angle, α:

$$I_{T(AV)} = \frac{V_m}{2\pi R_L}(1 + \cos\ \alpha)^* \qquad (28\text{-}14)$$

Thus, if $\alpha = 0°$, (conduction angle $\phi = 180°$), $I_{T(AV)} = \frac{V_m}{\pi R_L}$, as in an ordinary half-wave rectifier. If $\alpha = \phi = 90°$, $I_{T(AV)} = \frac{V_m}{2\pi R_L}$.

A smaller conduction angle will naturally produce a smaller average current.

The current rating of an SCR is usually given in terms of its average forward current rating, $I_{T(AV)}$, with the conduction angle equal to 180°, and at some specified case temperature T_C. For example, the 2N3668-2N3670 SCRs referred to earlier are rated at $I_{T(AV)} = 8$ A at a conduction angle of 180° and $T_C = 80°C$. If the device is used in a circuit with smaller conduction angles or at higher case temperatures, the SCR must be derated, as shown in Fig. 28-9.

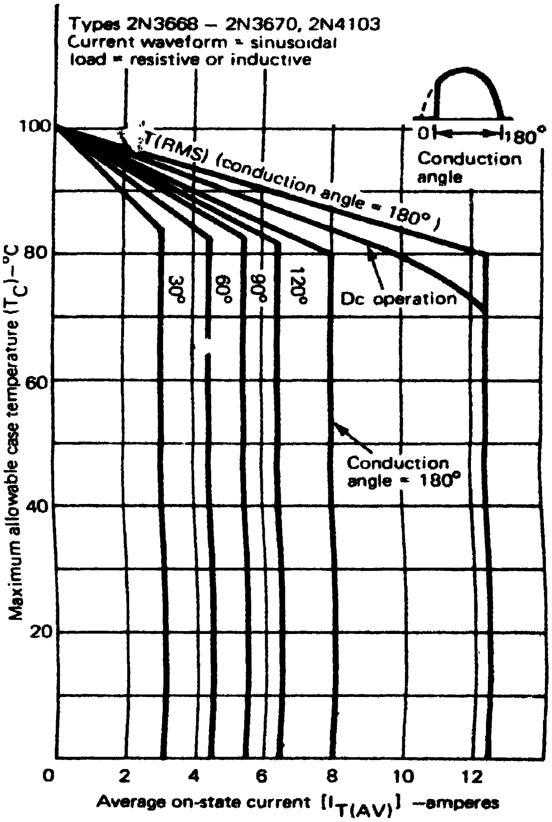


Fig. 28-9. Rating chart for type 2N3668-2N3670 SCRs showing how they must be derated for elevated case temperatures and smaller conduction angles. (Courtesy of RCA Corp.)

***See Millman, J., and Halkias, C., *Integrated Electronics*, McGraw-Hill, New York, 1972, Section 18-15.**

The need for derating at lower conduction angles should be obvious. If the device were operated at an average current of 8 A when the conduction angle is only 30° for example (instead of 180°), much higher peak values of current would be necessary. This could cause the average temperature of the SCR to be excessive and exceed the maximum operating junction temperature, which is typically 100°C.

The maximum average forward power dissipation of the SCR may be determined as for power transistors, using the following equation:

$$P_{AVG(max)} = \frac{T_{J(max)} - T_{C(max)}}{\theta_{J\text{-}C}}. \qquad (28\text{-}15)$$

For the 2N3668, $\theta_{J\text{-}C} = 1.7°C/W$, a typical value.

The derating chart in Fig. 28-9 shows a curve for $I_{T(RMS)}$, also at a conduction angle of 180°, which indicates a higher current of 12.5 A for case temperature operation up to 80°C. In fact, the 2N3668 is rated as a 12.5 A, 100 V SCR, the current rating understood to be in RMS amps. This is actually an inflated value for the SCR, since it only operates in a unidirectional rectifying mode. That is, if operating at a 180° conduction angle, at case temperatures below 80°C, the maximum permissible reading of a series connected dc ammeter is 8 A; an ac ammeter that reads true rms current would, at the same time, indicate 12.5 A for the same conditions. Common practice is to use the RMS value in specifying the rated current of an SCR. Typical current ratings for low, medium, and high current SCRs are as follows:

Low: $I_{T(RMS)}$ = 0.8 to 3.6 A,
V_{RRM} = 100 to 400 V.

Medium: $I_{T(RMS)}$ = 7.4 to 35 A,
V_{RRM} = 50 to 800 V.

High: $I_{T(RMS)}$ = 55 to 1000 A,
V_{RRM} = 50 to 1200 V.

Also usually specified is the peak surge (non-repetitive) on-state current, I_{TSM}. For example, a C290/291C high current SCR, with $I_{T(RMS)}$ = 470 A at 65°C case temperature, has I_{TSM} = 5500 A.

28-2 THE TRIAC

Construction

Many applications arise in which SCRs must be connected in anti-parallel. In this connection, with suitable triggering circuitry for each gate, both the positive and negative halves of a sine wave may be controlled in conduction. This is obviously not a rectifying application, so that the *one* device that can perform this function, the triac, makes no mention of rectification in its name.

The junction diagram shown in Fig. 28-10(a) indicates that a triac is a five-layer device. The dotted lines indicate how each half may be considered as a PNPN SCR with a common gate, as shown in Fig. 28-10(b).

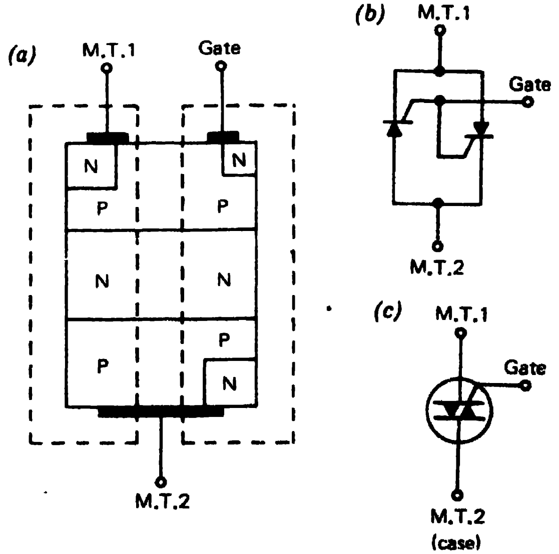


Fig. 28-10. (a) A cross-sectional diagram of a triac. (b) Showing how the triac resembles two SCRs in anti-parallel. (c) Standard symbol for the triac.

Note that the gate is closer to main terminal No. 1. (The terminology of anode and cathode is no longer used, since main-terminals No. 1 and No. 2 may be either anode or cathode depending upon the applied voltage.)

Characteristics

The principal V-I characteristics (the voltage and current between the main terminals) in Fig. 28-11 show that the triac is similar to the SCR in quadrant I, but, as expected, show a symmetrical characteristic in the third quadrant. That is, the triac may be triggered into conduction in either of the two quadrants *and* with either a positive or negative gate voltage. The four triac triggering modes are shown in Table 28-1.

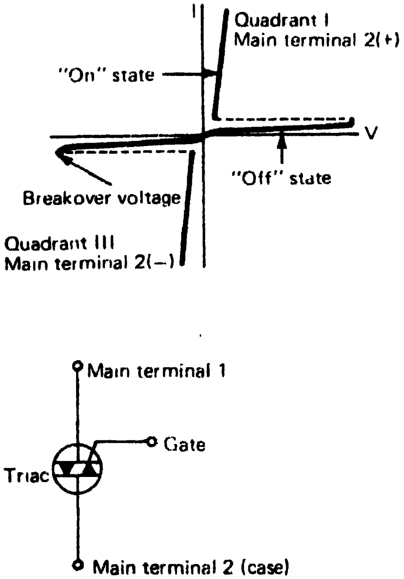


Fig. 28-11. Principal V-I characteristics of the triac.

Normal SCR operation occurs when main terminal No. 2 is positive with respect to main terminal No. 1 and the gate is positive with respect to main terminal No. 1. This is referred to as operating quadrant I(+). In this mode, the gate is the most sensitive in triggering the triac on. In the other modes, more current is needed.

TABLE 28-1 TRIAC TRIGGERING MODES

GATE-TO-MAIN-TERMINAL NO 1 VOLTAGE	MAIN-TERMINAL-NO 2-TO MAIN TERMINAL-NO 1 VOLTAGE	OPERATING QUADRANT
Positive	Positive	I (+)
Negative	Positive	I (−)
Positive	Negative	III (+)
Negative	Negative	III (−)

For example, an 8 A 200 V 40668 triac requires the following gate currents: I (+), 10 mA; III (−), 15 mA; I (−), 20 mA; III (+), 30 mA.

All terminology and parameters used with the SCR apply equally well to the triac, with the obvious exception of the reverse voltage capability. Two other differences are to be noted.

First, triacs are not available in as high a current rating as the SCR. The maximum available in 1972 was 80 A at a V_{DROM} of 600 V. Second, the current rating is given in RMS amps ($I_{T(RMS)}$) for a 360° conduction angle at a specified case temperature. This is to be expected, since the triac is used in applications where conduction occurs on both positive and negative half cycles, and average current would be meaningless. As with the SCR, the RMS conduction current capability must be reduced at elevated case temperatures.

28-3 TRIGGERING DEVICES

Although the triac may be fired into the conducting state by a simple resistive triggering circuit, more reliable and faster turn-on may be had if a switching device is used in series with the gate. As indicated in Fig. 28-1, triggering devices may be of the unidirectional or bidirectional type.

The Silicon Unilateral Switch (SUS)

This is a four-layer diode that resembles the PNPN SCR, in which contact is made only to the anode and cathode. Its V-I characteristics are similar, with the forward switch-on voltage in the region of 6 to 10 V. However, its maximum forward current capability is only in the order of 200 mA. Its symbol is shown in Fig. 28-12(a), along with the symbol for a Shockley diode; this is similar to the SUS, but its forward voltage drop after firing is lower.

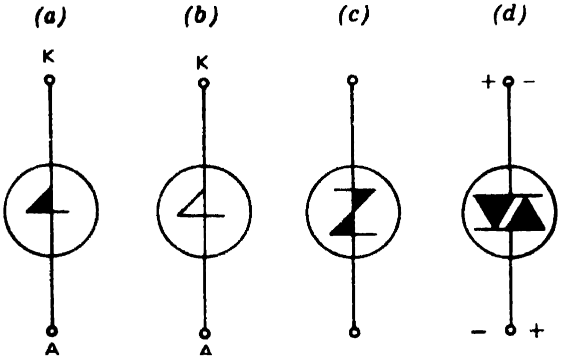


Fig. 28-12. Symbols for triggering diodes. (a) Silicon unilateral switch (SUS). (b) Shockley diode. (c) Silicon bilateral switch (SBS). (d) Diac.

The Silicon Bilateral Switch (SBS)

This is a bidirectional triggering diode resembling the triac construction but without a gate contact. Its characteristics are symmetrical, breaking down and switching on for an applied voltage of either polarity which is typically between 6 and 10 V. Its symbol is shown in Fig. 28-12(c). Its current capability is also 0.2 A.

The Diac

Specifically designed to supply the large gate current for fast turn-on of SCRs, the diac is a three-layer device resembling a bipolar PNP transistor. It is different in that the doping concentrations at the two junctions are approximately the same. This gives a symmetrical bidirectional switching characteristic with a forward (or reverse) breakover voltage, V_{BO}, between 29 V and 35 V for a 1N5411. It can handle a peak pulse current of 2 A (for a pulse duration of 30 μsec).

Since the diac is bidirectional in nature, it is ideally suited for triggering triacs, as we will now show.

28-4 TYPICAL DIAC - TRIAC PHASE - CONTROL CIRCUIT

Fig. 28-13(a) shows a typical phase-control circuit that may be used for lamp dimming, heat control, or universal motor speed control. This is similar to the circuit in Fig. 28-7(a), with the diac replacing D1 and an extra time constant, R_3C_3, added. This extends the effective range of R_2 in providing light control, especially at small conduction angles.

The operation of the circuit is similar to that described earlier for the SCR. As the input voltage increases, either positively or negatively, C_2 and C_3 charge at a rate determined primarily by R_2. When the voltage across C_3 exceeds the breakover voltage of the diac, the diac conducts. The capacitor discharges through the conducting diac into the gate of the triac, and the triac is triggered on. The breakover voltage of the diac, in the region of 30 V, ensures a gate current pulse of sufficient magnitude and duration to provide fast turn-on of the triac. A more closely controlled switching point is obtained with a diac than with simple resistance-capacitance triggering circuits, since gate characteristics can vary widely from one unit to another and also vary with temperature. The diac is less temperature-sensitive and has closer controlled characteristics.

As seen by the waveforms in Fig. 28-13(b), the triac will also be triggered into conduction

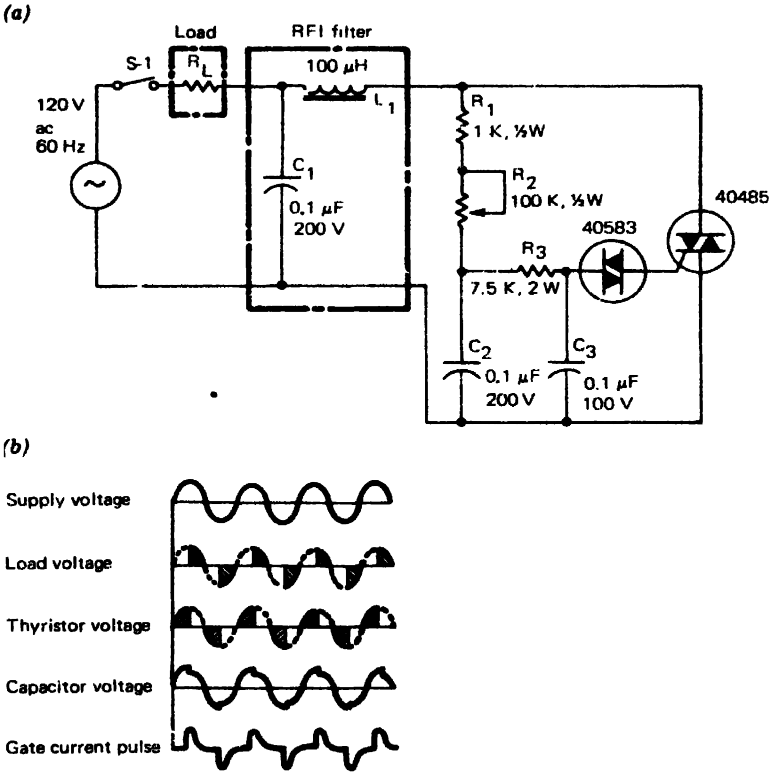

(a)
Load
RFI filter
S-1
R_L
100 μH
L_1
120 V
ac
60 Hz
C_1
0.1 μF
200 V
R_1
1 K, ½W
R_2
100 K, ½W
40485
40583
R_3
7.5 K, 2 W
C_2
0.1 μF
200 V
C_3
0.1 μF
100 V
(b)
Supply voltage
Load voltage
Thyristor voltage
Capacitor voltage
Gate current pulse

Fig. 28-13. (a) Schematic of a double-time-constant phase-control circuit suitable for light dimming, heat control, and universal motor speed control. (b) Waveforms of voltage and current in the diac-triac control circuit.

on the negative half cycle. Adjustment of R_1 can provide continuous control of load voltage from practically zero to practically full RMS value.

Radio-Frequency Interference (RFI)

The waveform of load voltage in Fig. 28-13(b) indicates that current rises very rapidly (in 1 or 2 μsec in a resistive circuit) each time the triac conducts. This rapid rise produces noise or radio frequency interference (RFI), which may be radiated and can cause direct interference in AM radios in the *immediate* vicinity. However, the interference that is more objectionable is that carried through the power lines to equipment operated from the same lines. Two methods may be used to reduce RFI, according to the application.

In light-dimming and speed controls, where the triac must be fired at least once each cycle to avoid flicker or "rough" motor operation, an RFI filter may be used as indicated in Fig. 28-13(a). The values shown for L_1 and C_1 are adequate for most light-dimming applications.

A second method of avoiding interference may be used in heating controls, in which the current through an electrical heater may be turned on for a few cycles then turned off for a few cycles, to maintain a given desired temperature. If the turn-on is only allowed to take place when the voltage is zero, no rapid increase in current takes place, as it does if switching occurs *during* the cycle.

A silicon monolithic integrated circuit that provides this function is the CA 3059 zero-voltage switch. As the name implies, this circuit provides an output pulse (typically 100 μsec duration) each time the applied line voltage crosses zero, on each half-cycle. The pulses are suitable for triggering SCRs or triacs for frequencies from 50 Hz to 400 Hz. When a ramp voltage generator circuit is connected to the zero-voltage switch, the interval during which these pulses are applied to a thyristor may be varied, which controls the average current and heat.

This circuit may also be used with light-dimming circuits if the ac frequency is 400 Hz, as is found in aircraft applications. This higher frequency makes it possible to eliminate full or half cycles within the period of the ramp voltage (typically 17.5 msec) without producing noticeable flicker.

28-5 THE UNIJUNCTION TRANSISTOR (UJT)

Another triggering device that may be used to fire thyristors is the unijunction transistor. Like the triggering diodes covered earlier, the UJT has a negative-resistance region. This is used to advantage in a relaxation oscillator whose frequency may easily be varied to provide an excellent method of motor speed control.

Construction

Although modern methods of automatic production result in a UJT that is annular in nature (see Fig. 28-14(b)) the earlier bar type structure shown in Fig. 28-14(a) helps to explain the symbol of the UJT and its operation. Basically, it consists of a bar of N-type silicon with one P-type emitter providing one PN junction. Note that the emitter is closer to base-two than base-one. This is clearly shown in the symbol in Fig. 28-14(c).

Static Characteristic Curves

Typical emitter characteristic curves for a UJT are shown in Fig. 28-15. If the voltage from base-two to base-one, V_{BB}, is made zero, and a voltage is applied to V_E, the resulting current I_E that flows gives the emitter to base-one diode characteristic. With $V_{BB} = 10$ V or 20 V, there is a leakage current I_{EO} from base-two to the emitter. It takes approximately 7 volts from emitter to base-one to reduce this current to zero and then cause current to flow in the opposite direction, reaching a peak point, V_P. After this point, an attempt to increase V_E is followed by a sudden increase in emitter current I_E, with a corresponding decrease in V_E. This is the unstable, negative-resistance portion of the characteristics, which lasts until the valley voltage is reached and the device saturates.

Equivalent Circuit

The equivalent circuit in Fig. 28-16 is quite straightforward. The resistance of the silicon bar is represented by two series resistors. R_{B2} is the resistance of the base-two portion, while R_{B1} is shown as a variable resistance, its value depending upon the bias voltage, V_D. The one PN junction (unijunction) is represented in the emitter by a diode.

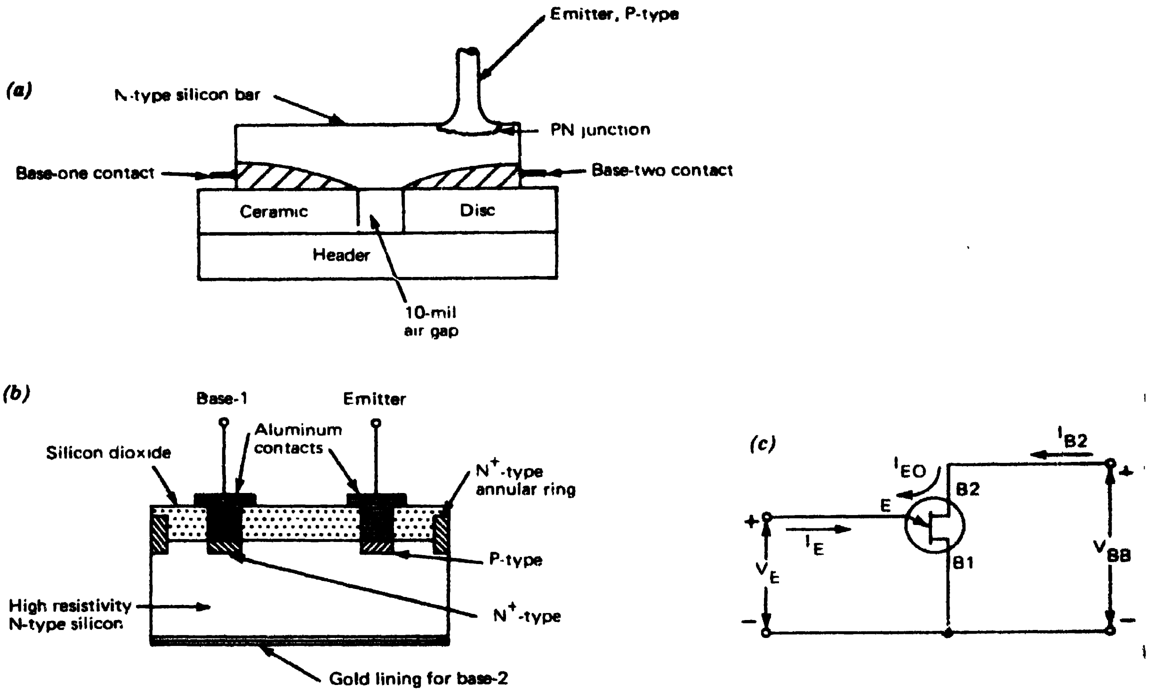

(a)
Emitter, P-type
N-type silicon bar
PN junction
Base-one contact
Base-two contact
Ceramic
Disc
Header
10-mil air gap
(b)
Base-1
Emitter
Aluminum contacts
Silicon dioxide
N^+-type annular ring
P-type
High resistivity N-type silicon
N^+-type
Gold lining for base-2
(c)
I_{B2}
I_{EO}
B2
E
+
I_E
V_E
B1
V_{BB}
−

Fig. 28-14. (a) Bar type of UJT structure. (b) Annular structure of a unijunction transistor. (c) Symbol of the UJT. (Courtesy of General Electric Semiconductor Products Department, Syracuse, N.Y.)

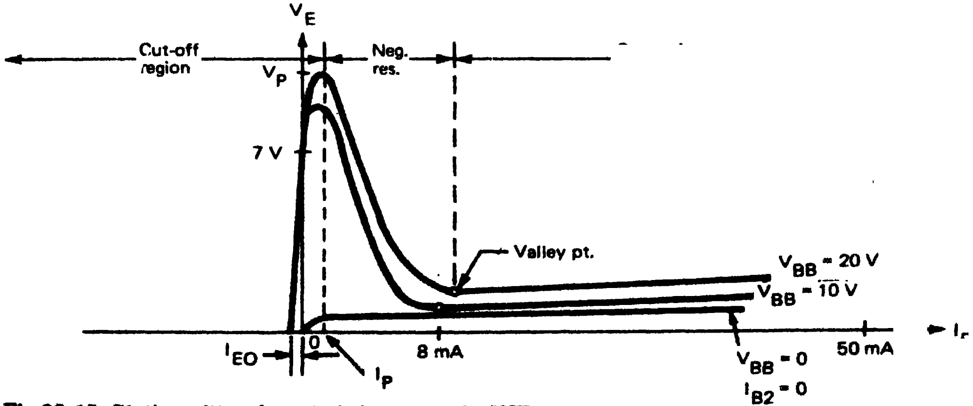


Fig. 28-15. Static emitter characteristic curves of a UJT.

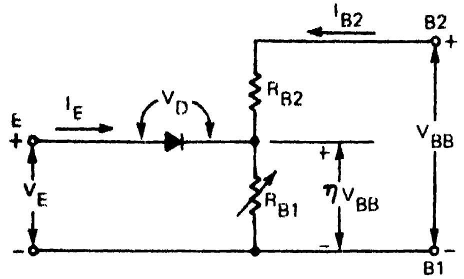


Fig. 28-16. Equivalent circuit of a UJT showing the single PN junction and the interbase resistances.

With no voltage applied to the UJT, the interbase resistance can be measured with an ohmmeter, such that

$$R_{BB} = R_{B1} + R_{B2}. \tag{28-16}$$

This is typically 4.7 K to 9.1 K for a 2N1671.

If a voltage is applied at V_{BB} with E open, the voltage will divide up across R_{B1} and R_{B2}, so that the fraction of the applied voltage appearing across R_{B1} is given by

$$\eta V_{BB} = \frac{R_{B1}}{R_{B1} + R_{B2}} \times V_{BB}, \tag{28-17}$$

where η = intrinsic stand-off ratio

$$\frac{R_{B1}}{R_{B1} + R_{B2}} \tag{28-18}$$

and is typically between 0.51 and 0.82.

Thus, the diode is reverse biased. Now with a fixed voltage at V_{BB}, say 10 V, consider increasing V_E until V_E just exceeds ηV_{BB}. The diode will become forward biased, and holes will be injected into the silicon bar, with a resulting decrease in resistance from emitter to base-one. Therefore, for an applied I_E, with the resistance decreasing the voltage V_E must decrease, giving a negative-resistance effect, in which current increases with decreasing voltage.

At the peak voltage,

$$V_P = \eta V_{BB} + V_D, \tag{28-19}$$

where V_D is typically 0.5 V.

Thus, if V_P is determined experimentally for a given V_{BB}, η can be determined for the given transistor. Knowing R_{BB} from a measurement will permit finding the values of R_{B1} and R_{B2}. The variation in R_{B1} with emitter current I_E can be seen from Table 28-2. The way in which V_E drops with increasing I_E is also evident.

TABLE 28 2 SHOWING THE VARIATION IN R_{B1} WITH NOTE HOW V_E DECREASES, THEN INCREASES.

I_E, mA	R_{B1}, ohms	$V_E \approx I_E R_{B1}$
0	4600	0
1	2000	2 V
2	900	1.8 V
5	240	1.2 V
10	150	1.5 V
20	90	1.8 V
50	40	2.0 V

When the diode becomes forward biased, there is an increase in I_{B2}, therefore a current limiting resistor R_2 is placed *external* to the transistor in series with base-two. The value of R_2 is given as

$$R_2 \approx \frac{1000}{\eta V_{BB}} \tag{28-20}$$

for the 2N2646 and 2N2647 unijunction transistors.

$$R_2 \approx \frac{0.4\, R_{BB}}{\eta V_{BB}} + \frac{(1 - \eta) R_1}{\eta} \tag{28-21}$$

for the 2N489 series, 2N1671, and 2N2160 unijunction transistors.

Typical values of R_2 are 100 to 200 ohms. A resistor is placed in the base-one lead to develop the output voltage; usually R_1 is 39 to 47 ohms. This can best be seen in the application of the UJT in an oscillator circuit.

Application of the UJT as a Relaxation Oscillator

The negative-resistance portion of the characteristics is used to good advantage in a relaxation oscillator where the discharging of a capacitor through the UJT can develop a pulse of output voltage. See Fig. 28-17.

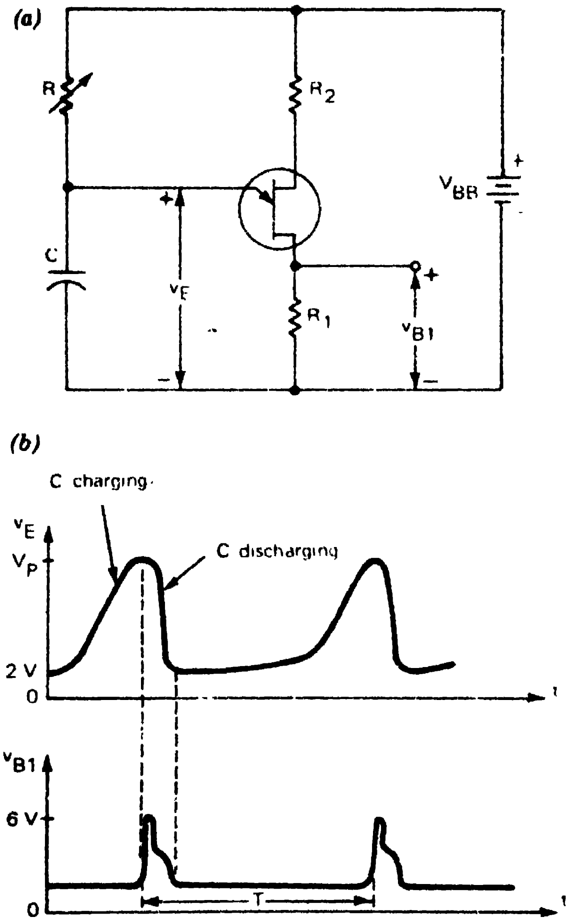


Fig. 28-17. (a) Circuit of a UJT relaxation oscillator. (b) Voltage waveforms.

C charges through R from V_{BB} until v_E reaches V_P and the UJT turns on. C discharges rapidly through R_1, giving a pulse of output voltage v_{B1}. When v_E drops to about 2 V, the emitter ceases to conduct, and the UJT switches off, allowing C to charge again.

The frequency of oscillation depends upon the time constant RC and the transistor, but the frequency is fairly independent of V_{BB} and temperature.

The period of oscillation is given by

$$T = \frac{1}{f} = RC \ln \frac{1}{1-\eta} = 2.3\, RC \log_{10} \frac{1}{1-\eta}. \qquad (28\text{-}22)$$

For a given η, this can be calculated, but it can also be obtained from a nomogram. This is a diagram in which two lines are drawn for R and C, and, assuming a certain average value of η, a third line indicating frequency and period is drawn. The intersection of this line for given values of R and C determines the frequency, as shown in Fig. 28-18, which has been drawn for $\eta = 0.62$.

If R is made variable (from 3 K ohms to 3 M ohms), an oscillator whose frequency can be varied results. Although such an oscillator using a dc supply voltage could be used to trigger a thyristor, there would be trouble synchronizing the pulses with the cycles of the alternating current. Fig. 28-19 illustrates a stable triggering circuit for an SCR in which the firing angle can be varied from 0° to 180°.

The low output impedance of the UJT (39 ohms) is ideal for driving the SCR, which has a relatively low input impedance from gate to cathode. Resistor R_D is used as a dropping resistor to restrict the peak voltage across the UJT to within its specifications.

The UJT may also be used in precision timing-delay and current-sensing circuits. The reader is referred to texts on industrial electronics for these applications.

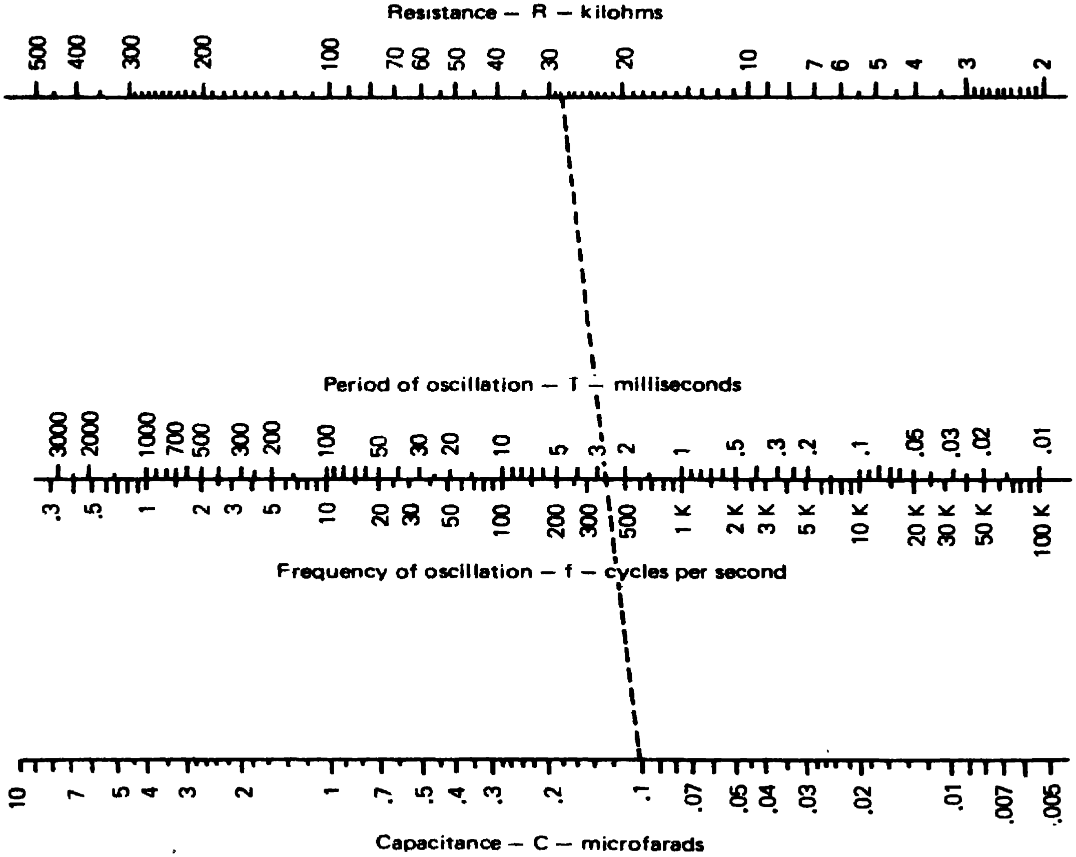
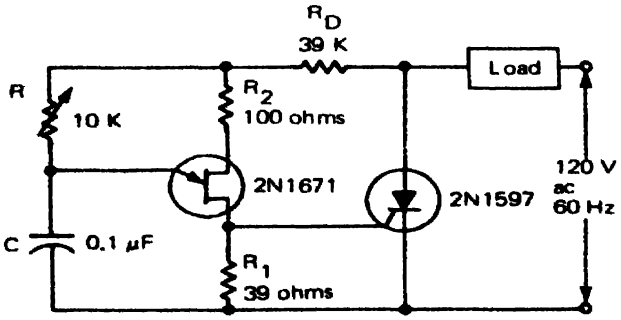


Fig. 28-18. Nomogram for determining frequency or period of a UJT relaxation oscillator. (Courtesy of General Electric Semiconductor Products Department, Syracuse, N.Y.)


R_D
39 K
Load
R
10 K
R_2
100 ohms
2N1671
2N1597
120 V
ac
60 Hz
C
0.1 μF
R_1
39 ohms

ig. 28-19. UJT relaxation oscillator synchronized ith the ac line to provide stable triggering of the CR.

28-6. SUMMARY

(1) Thyristors are solid-state PNPN switching devices which include the unidirectional SCR and the bidirectional triac.

(2) The SCR may be triggered into a conducting state by a pulse of gate current, but the gate loses control of anode current after firing.

(3) An SCR may be turned off by reducing its current below the holding value or by reducing the anode-to-cathode voltage to zero.

(4) The SCR may be thought of as a PNP and an NPN transistor interconnected, with regenerative feedback that is initiated by a gate current pulse.

(5) Gate triggering of a thyristor requires the application of gate voltage and current that must lie in the preferred area of the graphical gate characteristics to ensure reliable switching.

(6) Simple variable half-wave and full-wave rectifiers employing an SCR permit dc power control by varying the conduction angle using various trigger circuits.

(7) The current rating of an SCR may be given in average amps ($I_{T(AV)}$) or RMS amps ($I_{T(RMS)}$) for a conduction angle of 180° and a given case temperature. These must be derated according to the conduction angle and case temperature. Triacs are rated only in RMS amps.

(8) A triac is a bidirectional thyristor, conducting in either direction for either a positive or negative gate voltage. It has four triggering modes, as summarized in Table 28-1.

(9) Many triggering devices exist to fire thyristors, but the diac, a three-layer PNP diode, is most often used, since it can handle a peak current of 2 A for 30 μsec to promote fast turn-on.

(10) A typical diac-triac phase control circuit may employ one or two RC time constants in the triggering circuit. In this way, current through a series-connected load may be varied from 0° to 360° conduction in light, heat, and motor speed control applications.

(11) The rapid rate of rise of current in switched thyristor circuits can cause radio frequency interference (RFI). In light dimming applications, a simple LC filter circuit is sufficient to suppress the noise generated, but in heat controls a zero-voltage switching technique may be employed.

(12) The silicon unijunction transistor exhibits a negative resistance effect that is used to advantage in a relaxation oscillator circuit. This circuit is ideally suited for firing thyristors at a variable rate.

QUESTIONS

Q28-1. (a) What is a unidirectional thyristor?
 (b) How does an SCS differ from an SCR?

Q28-2. (a) In what ways can an SCR be triggered on?
 (b) What is the normal way of making an SCR conduct?
 (c) What is meant by the *latching current*?
 (d) What is the significance of the *holding current*?
 (e) What is a typical voltage drop across a conducting SCR?

Q28-3. (a) In the two-transistor analysis of an SCR, what does it mean for $G = h_{FE1} \times h_{FE2}$ to approach unity (one)?
 (b) Why would it be impossible to construct a germanium SCR? (Consider G near zero and being made to approach one.)

Q28-4. (a) What does the preferred area of gate triggering consist of in a thyristor? Explain fully the boundaries.

(b) How does the necessary gate triggering current required for a given device vary with temperature? Why?
(c) If a thyristor will switch on with $I_{GT} = 50$ mA, why use a current pulse of 200 mA?
(d) What problems can arise from excessively fast turn-on of a thyristor?

Q28-5. (a) What are the ranges of firing and conduction angles in a simple variable half-wave rectifier as in Fig. 28-6?
(b) Why can the firing angle not be extended beyond 90° in this circuit?
(c) What makes it possible to extend the firing angle in the circuit of Fig. 28-7?
(d) How often should gate control current be applied in Fig. 28-8, compared with the half-wave rectifiers, to obtain a controlled full-wave rectification?

Q28-6. (a) What two current ratings may be used in an SCR?
(b) Which is numerically larger for the same condition? Why? (Consider the mathematical definition of the two.)
(c) Why must the average current rating be decreased at smaller conduction angles than 180°?
(d) Which equation shows that heat-sinking a thyristor is just as important as in power transistors?

Q28-7. (a) Is the triac perfectly symmetrical in its construction?
(b) How is this supported as far as the triggering requirements are concerned in the different modes of firing?
(c) Why should the gate be most sensitive when triggering in the I (+) mode?
(d) How does the current rating of a triac differ from that of an SCR? Why?
(e) Why do you think that triacs are not available in as high a current rating as SCRs? (Consider the problem of changing from a conducting to a nonconducting state in a very short time with large conduction angles.)

Q28-8. (a) What advantage does a diac have over a silicon bilateral switch?
(b) What advantage could the SBS have?
(c) If an SUS were used to trigger a triac, what would be the effective operation of the circuit?
(d) Can a diac be used to trigger an SCR just as well as a triac?
(e) What advantage is provided by using a diac for triggering instead of a simple RC firing circuit?

Q28-9. (a) What is RFI?
(b) In what two ways does it manifest itself?
(c) Describe two methods by which RFI may be suppressed or eliminated.
(d) Why could the zero-voltage crossover switch not be used in light dimming applications at 60 Hz but could be used at 400 Hz or higher?

Q28-10. (a) What are the three main regions in a UJT's static characteristics?
(b) Which is used in an oscillator application?
(c) Define *interbase resistance* and *intrinsic stand-off ratio* as applied to a UJT.
(d) What must the emitter voltage increase to before the transistor switches from near cutoff to saturation?
(e) What three factors determine the frequency of oscillation of a UJT relaxation oscillator?

Q28-11. (a) Explain the operation of the circuit in Fig. 28-19.
(b) What is meant by *synchronizing* the oscillator?
(c) Is it possible for the oscillator frequency to be higher than 60 Hz?
(d) Is it possible that the oscillator frequency may be lower than 60 Hz?

(e) What effect would a higher frequency have?
(f) Could the circuit in Fig. 28-19 be used to trigger a triac? Could the triac be made to conduct for 360°?
(g) If the relaxation oscillator were driven from a separate full-wave rectified source (to provide synchronization), could the triac be controlled on both positive and negative alternations? What triggering modes would this constitute?

PROBLEMS

P28-1. A 2N3668 SCR is triggered from a dc source that has an open circuit voltage of 20 V.

(a) If the internal resistance of the source is 2 ohms, determine if reliable triggering is assured at a junction temperature of 25°C. (Hint: Draw the trigger load line one the characteristic curves of the 2N3668 in Fig. 28-5 as in Chapter 2, with $V_{GT} = 20$ V on one axis and $I_{GT} = V_{GT}/R_{INT}$ on the other axis as end points.)
(b) What are the two extreme conditions of gate voltage and gate current that might be encountered in this type of SCR with the given trigger source? What are the corresponding gate dissipations?
(c) Would this trigger source provide reliable triggering at –40°C? At +100°C?

P28-2. Repeat P28-1 using instead an internal resistance of 20 ohms.

P28-3. A 2N3668 SCR is triggered from a dc source of open circuit voltage 40 V. What maximum internal resistance may it have if reliable triggering is required at

(a) $T_J = 25°C$, (b) $T_J = 100°C$?

P28-4. A 2N3668 SCR is being triggered from a single-time-constant RC phase-shift circuit in which the voltage across the capacitor (and the gate of the SCR) increases to a peak of 40 V. If, at the time of firing, the circuit can be considered to have an internal resistance of 2 ohms, determine what minimum voltage on the capacitor will ensure switching at $T_J = 100°C$ and describe the *locus* of the gate triggering points as the voltage increases from zero. Is it possible that some units will trigger earlier than the voltage needed to *ensure* switching of *all* units?

P28-5. A 2N3668 SCR is triggered from a 6 V source of negligible internal resistance by supplying a pulse of current.

(a) If the SCR has the minimum gate resistance shown in the curves, what must the peak value of current be limited to, what must be the maximum duration of the pulse, and what is the peak gate dissipation for this condition?
(b) If the gate pulse duration is 10 msec, what must the amplitude of the current pulse be limited to?

P28-6. If a diac that conducts at 30 V is used between the capacitor and the SCR gate in P28-4, describe how the triggering action is different and which one will produce a more constant triggering point.

P28-7. A 200 A high current SCR is observed to have its anode current increase to 50 A in 0.4 μsec. What must be the minimum critical rate of rise of on-state current rating, di/dt, for this device?

P28-8. Consider the simple half-wave rectifier in Fig. 28-7.

(a) Sketch the input voltage waveform, v_i, for a 120 V 60 Hz input ($V_m = 170$ V).
(b) Sketch the waveform of v_C, in magnitude and proper phase relationship with v_i, using Eq. 28-13, for the following conditions:
(i) R adjusted to minimum, 1 K ohm.
(ii) R adjusted to maximum, 100 K ohms.

(iii) R adjusted to 5 K ohms total.
(c) Assuming the SCR fires at $v_C = 6$ V, indicate the approximate firing angle and conduction angle for each of the conditions in (b), showing the resultant load voltage waveform for each one.

P28-9. A 2N3668 has a current rating of $I_{T(AV)} = 8$ A at a 180° conduction angle and a case temperature of 80°C. $I_{T(RMS)} = 12.5$ A for the same conditions. Using the rating chart in Fig. 28-9, determine the maximum permissible current the device can handle for the following conditions:

(a) $\phi = 180°$, $T_C = 90°C$, $I_{T(AV)} = ?$
(b) $\alpha = 60°$, $T_C = 60°C$, $I_{T(AV)} = ?$
(c) $\phi = 30°$, $T_C = 90°C$, $I_{T(AV)} = ?$
(d) $\alpha = 120°$, $T_C = 100°C$, $I_{T(AV)} = ?$
(e) $\alpha = 0°$, $T_C = 90°C$, $I_{T(RMS)} = ?$

P28-10. For readers familiar with integral calculus, this problem shows why $I_{T(AV)}$ and $I_{T(RMS)}$ are 8 A and 12.5 A, respectively, for the 2N3668.

(a) Find the rms value of half a sine wave (180° conduction) in terms of I_m, using the following definition:

$$I_{RMS} \quad \sqrt{\frac{1}{2\pi}\int_0^{\pi}(I_m \sin\theta)^2 d\theta},$$

given $\sin^2\theta = \dfrac{1 - \cos 2\theta}{2}$.

(b) Find the average value of half a sine wave over a full cycle, in terms of I_m, using the following definition:

$$\text{dc} \quad I_{AV} = \frac{1}{2\pi}\int_0^{\pi} I_m \sin\theta \, d\theta.$$

(c) Determine what I_m must be if $I_{dc} = I_{AV} = 8$ A.
(d) Determine what the corresponding value of I_{RMS} is for this value of I_m.
(e) Determine the factor by which I_{RMS} is higher than I_{AV}.

P28-11. Consider a single-time-constant version of the diac-triac phase-control circuit in Fig. 28-13, in which R_3 and C_3 are omitted. Let $R_1 = 3.3$ K ohms, $R_2 = 250$ K ohms, $C_2 = 0.1\ \mu F$.

(a) Sketch the input ac waveform, v_i, for a 120 V 60 Hz supply.
(b) Sketch the waveform of the voltage across C_2, using Eq. 28-13 to obtain the magnitude and phase, for the following conditions:
(i) R_2 minimum. (ii) R_2 maximum.
(c) Determine the value of R_2 to provide a phase shift of 45°. Also determine the magnitude of the capacitor voltage and sketch as in (b).
(d) On the three curves in (b) and (c), draw two horizontal lines at ±30 V to indicate the assumed breakover voltage of the 40583 diac. Hence determine the approximate firing points of the triac and sketch the load voltage waveforms for the three conditions of R_2.

P28-12. The intrinsic standoff ratio for a UJT is determined to be 0.62. A measurement of its interbase resistance indicates 7 K ohms. What are the transistor's static values of R_{B1} and R_{B2}?

P28-13. A UJT with $\eta = 0.62$ is used in a relaxation oscillator circuit with R = 5 K ohms and C = 0.05 μF.

(a) Use the nomogram in Fig. 28-18 to determine the period and the frequency of oscillation.
(b) Verify this result by using Eq. 28-22.
(c) Determine what R must be changed to in order to obtain a frequency of oscillation of 60 Hz.

(d) If C is increased to 0.5 μF, what should be the value of R if the frequency is to be 60 Hz?

(e) If the UJT's value of η is higher than 0.62, is the value of f determined from the nomogram higher or lower than the actual value?

APPENDIX A

BASIC CHARACTERISTICS OF THE TRIODE AMPLIFIER

A-0. INTRODUCTION

The high vacuum triode has largely been replaced by the transistor. A few applications still remain where the tube is used at high frequencies where the transistor may not operate at sufficiently high power levels. Such an example is found in the Telstar satellite (1962), where the transmitting frequency is 4.17 GHz (4,170 MHz) and the receiving frequency is 6.39 GHz. For the final power amplifier, a traveling-wave tube (TWT) is used in the transmitter since semiconductor devices are not available to operate at the necessary power level. However, transistors are used both in the receiver and transmitter at lower frequencies and power levels. Another application for tubes is in the broadcast industry, where a final power amplifier may have to deliver as much as 50 kW and the tube (water cooled) has a dissipation capability of near 100 kW.

Also, many tubes are still being used in the radio and TV industry, although these will eventually (but not rapidly) be replaced by integrated circuits.

However, many of the techniques of analysis that apply to a tube circuit can be applied to a transistor circuit, and since the tube is a little easier to understand it provides a good vehicle for introducing many new ideas.

A-1. TRIODE CONSTRUCTION

The triode was evolved from the diode by the addition of a third element called the control grid, by means of which the flow of current through the tube could be controlled by a separate voltage source. The symbol for the triode tube is shown in Fig. A-1.

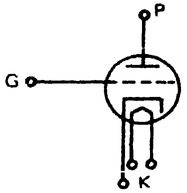


Fig. A-1. Symbol of a triode, showing the control grid G.

There are basically five types of tube construction. Typical tubes are shown in Table A-1 for each type. The *miniature* and *subminiature* types are relatively long compared with their diameter and may have a glass or metal envelope. Their grids consist of fine wires wound around vertical supporting rods with the cathode (and heater) on the inside and the plate on the outside. The grid is closer to the cathode than the plate.

In an attempt to compete with transistors a smaller more rugged type, called a *nuvistor*, was developed. This has a fine mesh for a grid and has all of its electrodes cylindrical in structure, with a metal shell.

For high frequency use, *planar ceramic* tubes were constructed that utilize ceramic for insulating the closely spaced electrodes and titanium for the anode. The grid is a flat mesh surface between the anode and cathode. Such tubes can work at frequencies as high as 7.5 GHz and rival the TO-5 transistor mounting in size.

Finally there is the *compactron* which, as its name implies, is quite compact, with its diameter and height very nearly equal.

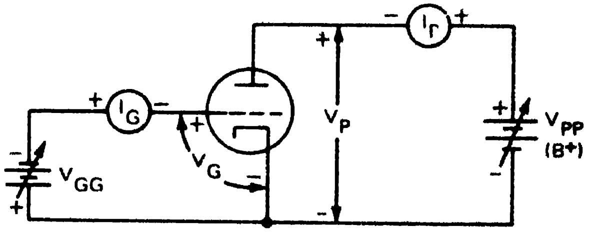
TABLE A-1 VARIOUS TYPES OF TRIODES USED AS AMPLIFIERS

TUBE NO	CLASSIFICATION	CONSTRUCTION	DIAM IN	SEATED HT IN	PLATE VOLTS	PLATE WATTS (PER SECTION)
5HA7	Dissimilar double triode	Compactron (Metal)	1.2	1.3	250	2.75,0.3
6FQ7/ 6CG7	Medium-mu twin-triode	Miniature (Glass)	0.75	2.4	250	4
6K4	Medium-mu triode	Subminiature (Metal)	0.4	1.5	200	3
12AE7	Double triode	Miniature (Glass)	0.75	2	12.6	1
7391	Hi-mu UHF triode (6 GHz)	Planar ceramic (Metal)	0.5	1	175	2
7720	Hi-mu VHF triode (500 MHz)	Planar ceramic (Metal)	0.33	0.5	250	1
7895	Hi-mu triode	Nuvistor (Metal)	0.4	0.6	110	1

Referring to Table A-1, it is seen that the 12AE7 operates with a plate voltage of only 12.6 volts. A whole line of such tubes is available for both audio and radio frequency use, suitable for operation in an automobile radio receiver without the need for a vibrator-power transformer-rectifier high voltage power supply.

A-2. TRIODE CHARACTERISTICS

The characteristics of a triode may be determined using the circuit in Fig. A-2. Note that capital letters are being used, since all voltages are dc and there is no time varying voltage applied. If $V_G = 0$, the tube behaves as an ordinary diode, so that, as V_{PP} is increased (numerically equal to V_P), the current through the tube increases, as shown by the curve for $V_G = 0$ V in the characteristic curves for the 6CG7, Fig. A-3(a).



V_P = plate to cathode volts
V_G = grid to cathode volts
I_P = plate current
I_G = grid current
V_{PP} = plate supply voltage
V_{GG} = grid bias supply voltage

Fig. A-2. Circuit to determine the characteristics of a triode.

When V_G is made -2 V, the presence of a negative voltage on the grid close to the cathode will reduce the electron flow across the tube (as the plate voltage is increased), below what it was when V_G was zero. The curve will be practically parallel, however,

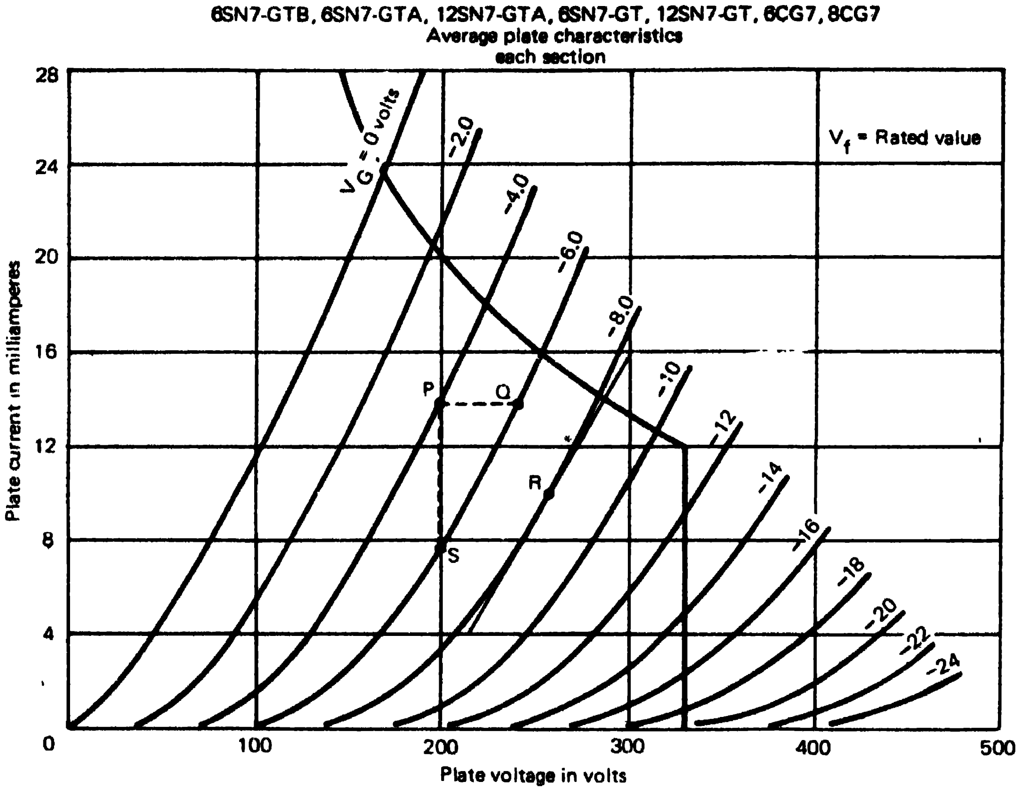


Fig. A-3(a). Average characteristic curves for the 6FQ7/6CG7, showing the permissible area of operation for the tube. (Courtesy of General Electric Co.)

to the first — just shifted down and to the right. As the control grid voltage is made more negative, less and less plate current flows for a given plate voltage. The resulting curves are referred to as the average plate characteristics.

A-3. TUBE PROPERTIES

There are three properties that are used to describe most high vacuum triodes, tetrodes, pentodes, etc.

Amplification Factor, μ (mu)

On the curves of Fig. A-3(a) consider points P and Q at which:

P : V_G = -4 V, V_P = 200 V I_P = 14 mA;

Q : V_G = -6 V, V_P = 242 V I_P = 14 mA.

This indicates that when V_G was changed from -4 V to -6 V the plate voltage V_P had to be increased from 200 V to 242 V to keep the current constant at $I_P = 14$ mA. (A change in V_G alone from -4 V to -6 V would have dropped the current to 8 mA. It required an increase in V_P of 42 V to bring the current back up to its initial value.) This is intended to show how much more control the grid has over the plate current than has the plate voltage. This is because the grid is located closer to the cathode than the plate and exerts more influence than the plate. The amount of increased control is measured by the *amplification factor*, μ. This is defined as follows:

$$\frac{\Delta V_P}{\Delta V_G} \text{ with } I_P \text{ held constant.} \qquad \text{(A-1)}$$

Using points P and Q:

$$\Delta V_P = 242 - 200 = 42 \text{ V};$$

$$\Delta V_G = -6 - (-4) = -6 + 4 = -2 \text{ V}.$$

$$\mu = \frac{42 \text{ volts}}{-2 \text{ volts}} = |-21| = 21.$$

Thus the grid of this tube, in this region, has 21 times as much control as the plate voltage over the plate current. This changes slightly depending exactly where on the characteristic curves it is evaluated, but over a large portion it is relatively constant at approximately 20.

The amplification factor should not be called the *gain* of the tube, although it is related to gain when used in an amplifier. Gain only has meaning in an amplifier circuit and may vary according to the type of the circuit, but the amplification factor is a property of the tube and is "constant" for a given tube. Actually, the μ of this tube does vary at low plate currents as can be seen in Fig. A-3(b), but over the region in which the tube is normally used the μ is constant. This is called a medium-mu tube. High-mu tubes have values around 100.

There are applications in which the mu of the tube is made variable purposely by making the grid around the cathode nonuniform. Then, by changing the grid voltage, the mu (and hence the gain of the amplifier in which the tube is being used) can be varied. This is used in automatic volume control (AVC) circuitry in radio receivers to reduce gain when strong signals are present.

Dynamic Plate Resistance, r_p

In the same way that the plate resistance of a diode can be found, so the plate resistance r_p of the triode can be determined. Refer to point R on the previous set of characteristic curves, Fig. A-3(a). A tangent has been drawn in which:

$$\Delta V_P = 300 - 218 = 82 \text{ V};$$

$$\Delta I_P = 16 - 4 = 12 \text{ mA}.$$

$$\text{Dynamic plate resistance, } r_p \quad \frac{\Delta V_P}{\Delta I_P}$$

$$\text{with } V_G \text{ constant.} \qquad \text{(A-2)}$$

$$r_p = \frac{82 \text{ volts}}{12 \text{ mA}} \qquad \frac{82}{12 \times 10^{-3}} = \frac{82 \times 10^3}{12} \text{ ohms}$$

$$= 6.83 \text{ K ohms.}$$

Again, r_p varies over the curves, but is approximately constant at 7 K ohms in this case, although values for other tubes may be 10 K ohms to 100 K ohms.

Mutual Conductance, or Transconductance, g_m

The mutual conductance involves holding V_P

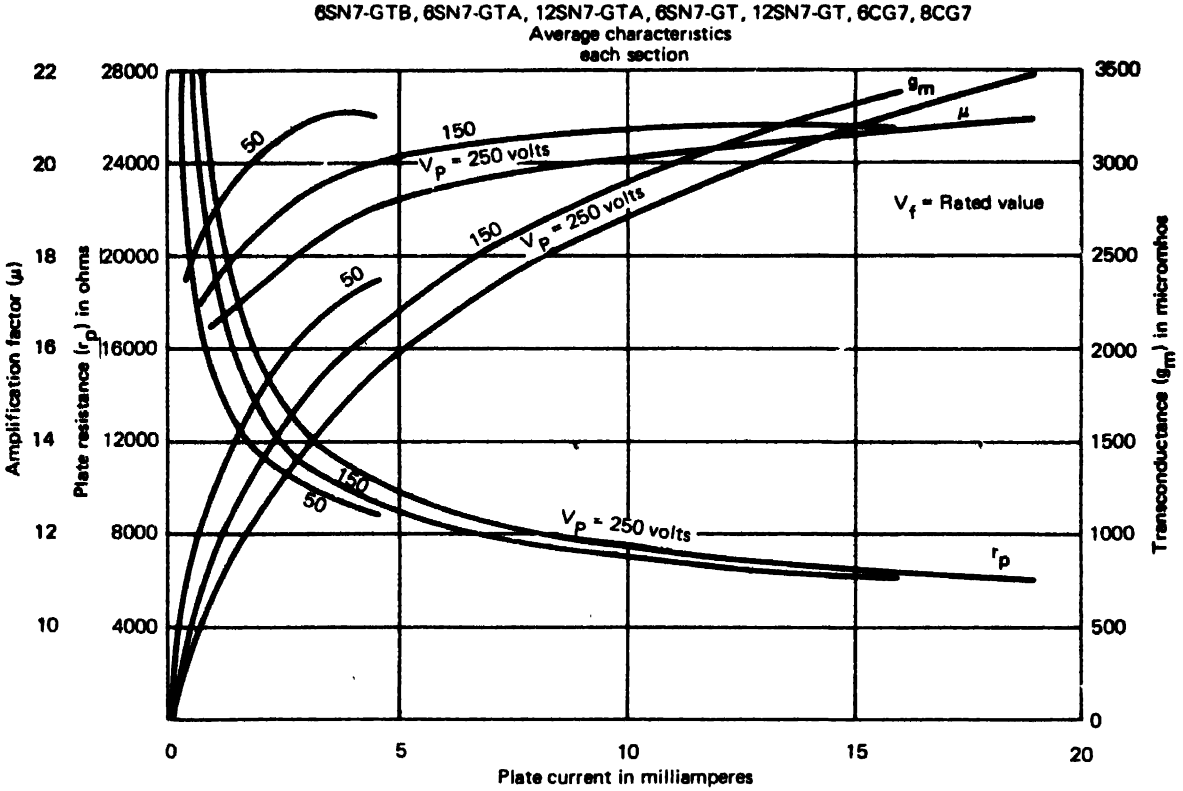


Fig. A-3(b). Average characteristics of the 6CG7 as a function of plate current. (Courtesy of General Electric Co.)

constant and taking the ratio of the change in I_p to the change in V_G:

$$g_m = \frac{\Delta I_P}{\Delta V_G} \text{ with } V_P \text{ constant.} \qquad \text{(A-3)}$$

For example, consider the point S on the characteristic curves and point P:

S: $V_G = -6$ V, $I_P = 8$ mA $V_P = 200$ V;

P: $V_G = -4$ V, $I_P = 14$ mA $V_P = 200$ V.

$\Delta I_P = 14 - 8 = 6$ mA.

$\Delta V_G = -4 - (-6) = -4 + 6 = 2$ V.

$$g_m = \frac{6 \text{ mA}}{2 \text{ volts}} = \frac{6 \times 10^{-3}}{2} \text{ mhos} = 3 \text{ millimhos}$$

$$= 3000 \text{ micromhos.}$$

g_m is subject to more variation over the curves than is μ or r_p.

Another set of curves can be drawn from the characteristic curves; these are called transfer characteristics, a few of which have been sketched in Fig. A-4.

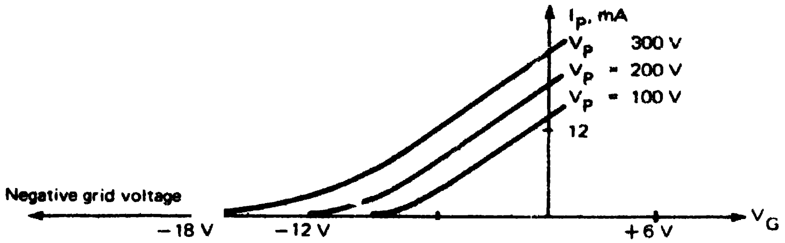


Fig. A-4. Transfer characteristics of a 6CG7 triode tube.

Since g_m is readily obtainable from these *trans*fer curves (since V_P = constant), g_m is more commonly called *trans*conductance. However, the same information is contained in either set of curves. The transfer curves help show whether a tube is a sharp cutoff or remote cutoff tube. If the curves become very flat for large negative grid voltages, it is a remote cutoff tube; if the curves are "sharp" for $I_P = 0$, the tube is a sharp cutoff type. The remote type is the one that has a variable mu and is used in AVC applications. A sharp cutoff tube may be used for switching purposes

A-4. RELATIONSHIP BETWEEN μ, r_p, g_m

Consider the ratio of μ to r_p:

$$\frac{\mu}{r_p} = \frac{\dfrac{\Delta V_P}{\Delta V_G}}{\dfrac{\Delta V_P}{\Delta I_P}} \qquad \frac{\Delta I_P}{\Delta V_G}$$

$$r_p g_m \qquad \text{(A-4)}$$

Note that r_p in ohms, multiplied by g_m in mhos, cancels to give a dimensionless number, μ, which is the ratio of volts to volts.

A-5. PERMISSIBLE AREA OF OPERATION

The separation between the electrodes of the tube places a limit on the maximum voltage that the tube can withstand. In addition, the physical size of the plate determines the amount of power and heat that it can safely dissipate. Consequently, there is a well-defined area of operation that the tube must be operated in and may be shown on the characteristic curves, as in Fig. A-3(a). The area is bounded by a vertical line corresponding to the maximum plate voltage (330 V for the 6FQ7/6CG7) and a curve called the maximum dissipation hyperbola, which is obtained as follows.

The product of plate voltage and plate current gives the plate dissipation, which for one section of the 6CG7 is 4 W. (The total for both sections is 5.7 W.) Thus for any selected voltage less than or equal to the maximum, there will be a corresponding maximum allowable plate current. For example, when V_P = 300 V, I_P is given by

$$I_P \quad \frac{\text{Power}}{300} \text{ amps} \quad \frac{4000}{300} \text{ mA}$$

$$= 13.3 \text{ mA.}$$

At V_P = 200 V, I_P = 20 mA, etc.

Plotting these points establishes the maximum dissipation hyperbola. Occasionally the maximum plate current that the tube can handle will provide an upper horizontal limit on this curve, but usually the tube is not operated in this region, so it is not an important factor except in pulsed operation.

A-6. OPERATION OF THE TRIODE AS A DC AMPLIFIER

Consider the circuit of a dc amplifier shown in Fig. A-5. The circuit is almost the same as Fig. A-2 except for the addition of a load resistor R_L, across which the output voltage is developed as a result of plate current that is controlled by the input signal voltage, V_s.

With reference to Chapter 7, we see that the source of energy that powers the amplifier is supplied by V_{PP} (also known as E_{bb} or the B+ supply), V_s with its internal resistance R_s supplies the signal, the tube is the active device, and in this example we have not yet connected any load R_o to the output.

Note that the signal to be amplified is dc with its negative terminal connected to the grid. Thus, if V_s is made equal to 4 V we refer to the input to the amplifier as $V_G = -4$ V. Voltages on the grid are understood to be with respect to the cathode, unless otherwise stated, since the characteristic curves are drawn for voltages on the grid with the cathode as the reference.

Example A-1. Let us assume we want to amplify the increase in V_s as it changes from -4 V to -2 V. Further, let it be given that the tube is a 6CG7 with characteristic curves as in Fig. A-6 and that V_{PP} is 300 V with R_L = 20 K ohms. Determine the voltage gain, A_v.

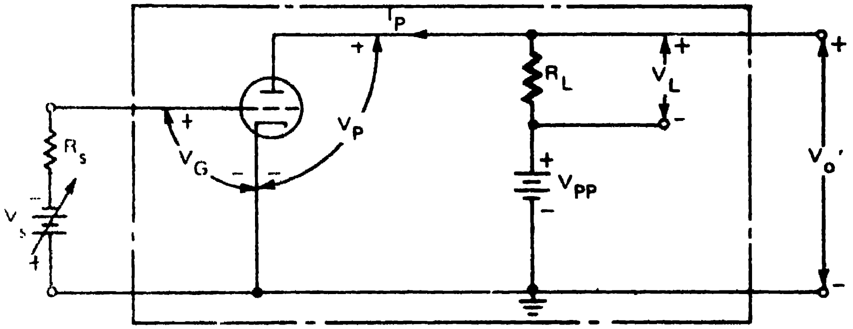


Fig. A-5. Basic circuit of a triode dc amplifier.

Solution. When $V_G = -4$ V, the curve relating the plate current, I_P, to the plate voltage (with respect to the cathode), V_P, is given by the third curve down on Fig. A-6. We are to determine the amount of current flow, V_L, and V_P (V_o') for this condition. Note that this condition corresponds exactly to that in Section 2-1, since the tube behaves like a modified diode with its characteristic curve shifted down and to the right compared with the curve obtained when the grid is grounded ($V_G = 0$ V). Consequently, we can draw a dc load line and find the values required from its intersection with the $V_G = -4$ V curve, using the following equation:

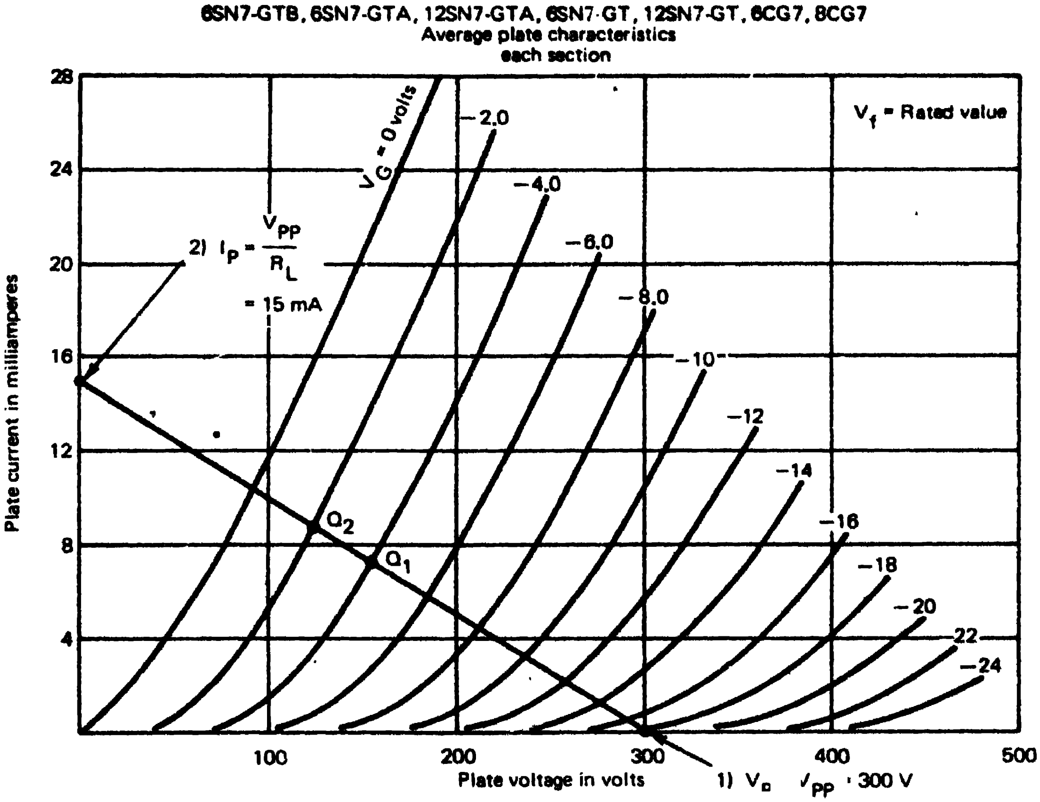


Fig. A-6. Dc load line drawn on characteristic curves for dc amplifier.

$$V_{PP} = I_P R_L + V_P. \qquad \text{(A-5)}$$

(1) Consider $I_P = 0$; then $V_P = V_{PP} = 300$ V.

(2) Consider $V_P = 0$; then $I_P = \frac{V_{PP}}{R_L} = \frac{300}{20\text{ K}}$

$$= 15 \text{ mA}.$$

This establishes the dc load line as shown in Fig. A-6, with the operating point at Q_1 (for V_G = -4 V) giving the following:

Q_1: $V_P = 155$ V, $I_P = 7.2$ mA.

Since V_o' is the same as the plate voltage V_P, the output voltage is 155 V dc. V_L is -145 V, since I_P makes the "bottom" of R_L positive with respect to the "top," which is opposite to the polarity markings used to describe V_L.

Now consider the input changing to -2 V. Since the grid is less negative, more plate current will flow, and more voltage will be dropped across R_L, with a resulting decrease in plate voltage. All of this is confirmed by examining the new quiescent point, Q_2, at the intersection of the dc load line and the V_G = -2 V curve. Note that a move *along the load line* from Q_1 to Q_2 corresponds to the change at the input of 2 V.

Q_2: $V_P = 125$ V, $I_P = 8.8$ mA, $V_L = -175$ V.

We may now determine the open circuit voltage gain of this dc amplifier by considering the ratio of the *change* in output voltage to the *change* in input voltage:

$$\frac{\Delta V_o'}{\Delta V_G} \quad \frac{\Delta V_P}{\Delta V_s} \quad \frac{155 - 125}{-4 - (-2)} \quad \frac{30}{-2} = -15.$$

The negative sign can be interpreted as an increase (numerically) at the input (from -4 V to -2 V) resulting in a *decrease* at the output (from 155 V to 125 V). When used with an ac amplifier it can be thought of as representing a 180° phase shift.

Note: We were not able to use $A_v = \frac{V_o'}{V_s}$ to determine gain, but instead had to consider the *change* in voltage levels, because this amplifier did not have a zero volt output with a zero volt input. In fact, if the input were zero ($V_s = V_g = 0$ V), the load line tells us that the output ($V_p = V_o'$) would be 92 V.

Observe then that a change in input voltage from –4 V to 0 V results in a change at the output from 155 V to 92 V, corresponding to a gain of –15.75. In other words, the voltage gain is fairly constant in this region of the curves.

Question: What is the input resistance of this amplifier?

Answer: Extremely high — theoretically infinite. Why is this? For the values considered above, the grid was either at a zero or negative potential, so that electrons did not flow in the grid circuit. As a result, no grid (or input) current meant that the input voltage to the amplifier V_G was always numerically the same as the source voltage V_s, regardless of the value of the internal resistance R_s. Because the signal source did not have to supply "any" current we term the tube amplifier a voltage-controlled device.

Question: What if the grid were driven positive?

Answer: This would depend upon the type of tube, the internal resistance of the source R_s, and the extent to which it was driven positive. Some tubes have grids rugged enough to dissipate heat and can withstand grid current, but this current means that the source must supply signal power. For sources with a high internal resistance, the result will be a distortion of the input signal (or reduction in value from V_s), during the portion of positive grid operation, since the input resistance to the amplifier will now be vastly reduced. In addition, some tubes show an unequal separation of the characteristic curves for equal increments of grid voltage when driven positive compared with when driven negative. This is especially true at low values of plate voltage, where grid current flows at the expense of plate current. Thus there are many advantages in not driving the grid positive. If the tube is operated so that the grid is always above cutoff (by avoiding highly negative grid voltages) and below positive grid operation, the amplifier is said to be operating in true class A. If, however, the grid is driven positive by any amount this is described as class A_2 operation, which may only be used with those tubes designed for this operation.

A-7. THEORY OF OPERATION OF AN AC AMPLIFIER; BIAS

If true class A operation (with all of its advantages) does not allow the grid to be driven positive, how is it possible to amplify a sinusoidally varying signal that varies positively and negatively from a zero reference? The answer lies in a bias supply voltage, which is a dc voltage placed in series with the ac signal to be amplified, as shown in Fig. A-7.

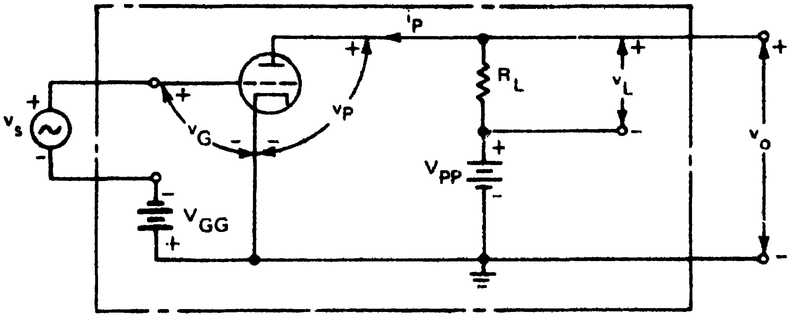


Fig. A-7. Basic circuit of an ac amplifier, showing the bias voltage V_{GG}.

V_{GG} is the bias voltage, such that when the signal voltage, v_s, is zero, the dc operating point of the tube (quiescent point) is placed somewhere in the middle of the characteristic curves, corresponding to a grid voltage of $V_G = -V_{GG}$. Now, as v_s increases positively, the bias voltage if properly selected will maintain the grid at a negative potential. That is, the minimum value of the size of the bias voltage must be equal to the peak of the input signal voltage, v_s. For example, if V_{sm} is +2 V, V_{GG} must be at least -2 V. If V_{sm} is +7 V, V_{GG} must be at least -7 V. Actually, the bias voltage could be made -4 V and -9 V respectively. It is not essential that the grid be driven to 0 V when v_s is at its maximum value. The actual value for V_{GG} will be determined from both the positive and negative swings in input voltage, v_s, so as to locate the operating point (with no input) in the middle of the characteristic curves. The variation in input voltage. v_s, will cause a similar variation in plate current developing an enlarged variation in voltage across R_L. The peak-to-peak output voltage, v_o', is equal in value to $V_{L,p\text{-}p}$ but opposite in phase to v_s, as shown by the following example.

A-8. GAIN OF AN AC AMPLIFIER

Example A-2. Determine the gain of the ac amplifier in Fig. A-7, and the waveforms of v_G, i_P, v_L and v_o', given the following: R_L = 20 K ohms, V_{PP} = 300 V, V_{GG} = 4 V, v_s is a sine wave whose maximum peak value is 4 V (peak-to-peak is 8 V).

Solution. In order to simplify the calculations, the characteristic curves have been slightly modified from those used in Example A-1 but represent the type of variation that may be encountered from tube to tube. The curves are shown in Fig. A-8 with the dc load line drawn as before. Note that it is still called a dc load line, even though we are analyzing an ac amplifier since we have not yet considered the addition of an output load resistor.

Now that we are working with an ac amplifier, we must be careful to use the appropriate small letters to indicate time varying voltages. Given a sine-wave and a square-wave input as shown in Fig. A-9(a), let us first determine the grid voltage waveform. The grid to cathode voltage is:

$v_G = v_s - V_{GG}$. Since V_{GG} is 4 V, v_G is the input sine wave with 4 V subtracted, which shifts the curve down so that its maximum value is 0 V and its minimum is -8 V. Effectively, it has been clamped at 0 V to avoid driving the grid positive, as shown in Fig. A-9(b).

The plate current that flows due to this

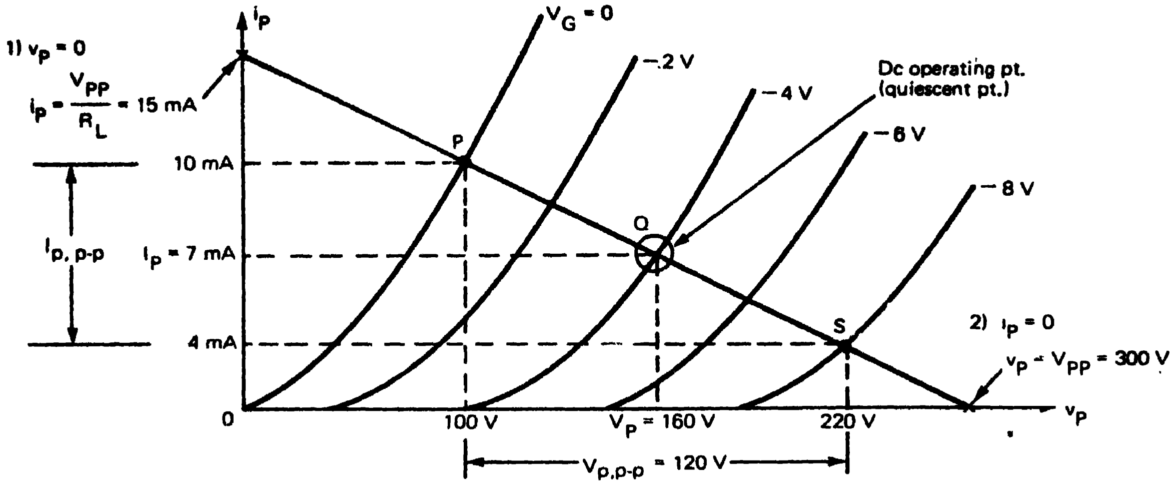


Fig. A-8. Dc load line drawn on characteristic curves for amplifier of Example A-2.

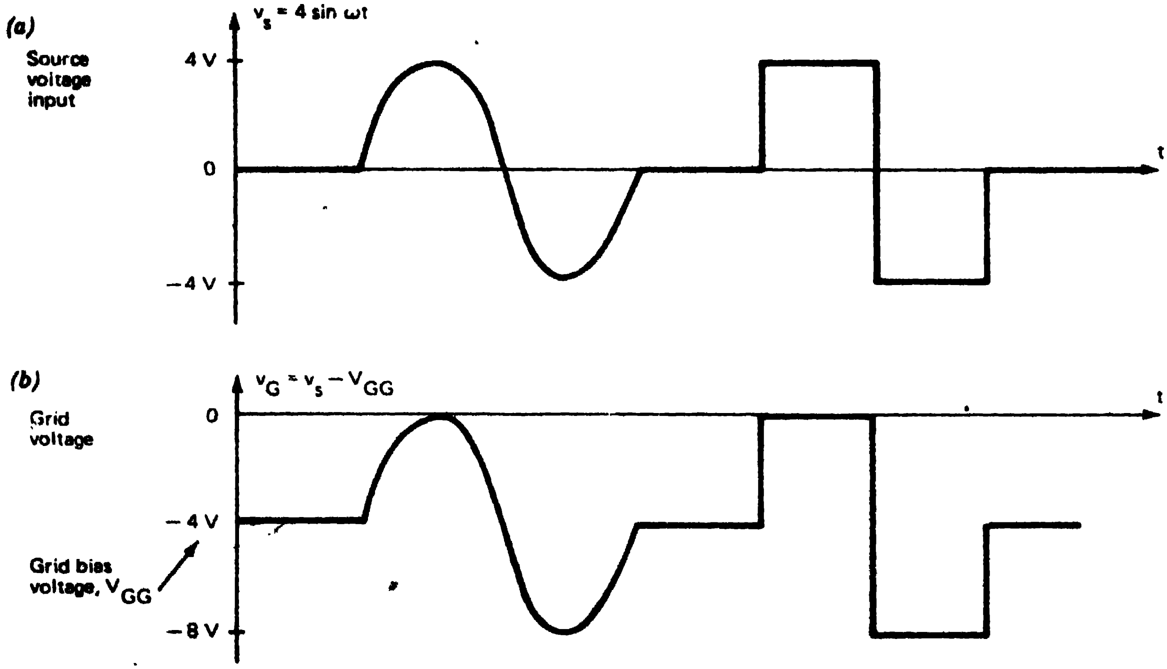


Fig. A-9. Waveforms in ac amplifier of Example A-2. (Sheet 1 of 2)

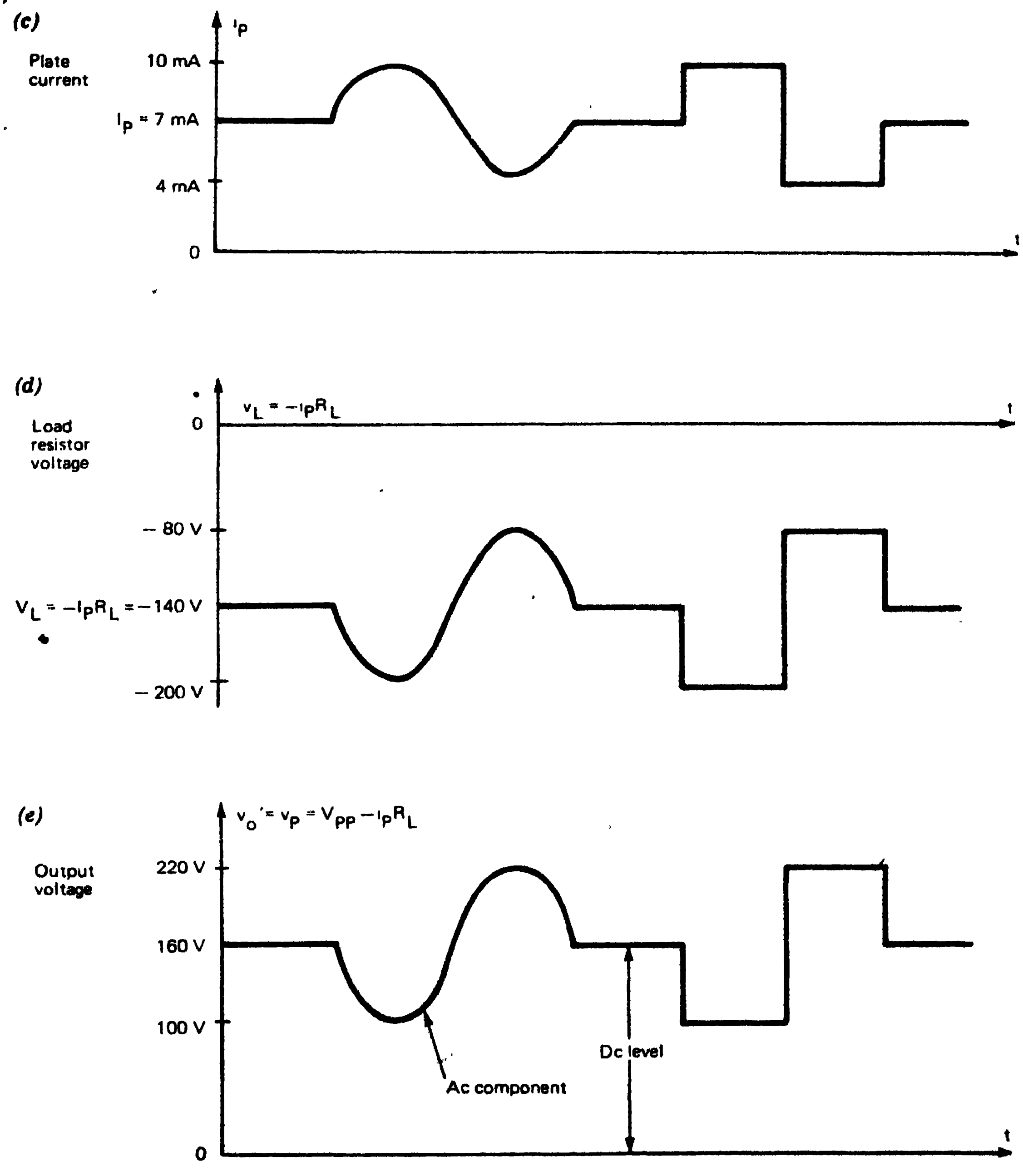


Fig. A-9. Waveforms in ac amplifier of Example A-2. (Sheet 2 of 2) grid voltage provides the voltage waveforms indicated in the rest of Fig. A-9. The steps involved to obtain these waveforms are summarized below.

(1) The intersection of the static load line with the bias voltage $V_{GG} = -4$ V gives the dc operating point, often called the Q-point or quiescent point. With $v_s = 0$, Q represents the static conditions in the tube, with $V_P = 160$ V, $I_P = 7$ mA.

(2) Now as v_s (a sine wave of 8 V peak-to-peak) is applied, v_G will change from -4 V toward 0 V. This is represented by moving *along the load line* from Q to P until a peak current $i_P = 10$ mA is produced, and v_P has dropped to 100 V at $v_G = 0$ V.

(3) As v_s decreases toward -4 V, v_G will change toward -8 V, *along the load line* through Q to S where $i_P = 4$ mA, $v_P = 220$ V.

(4) The output voltage v_o' can be read off directly from the v_P axis corresponding to P and S as 220 - 100 = 120 V peak-to-peak. The corresponding plate current variation is 10 - 4 = 6 mA peak-to-peak. Thus the plate current and volts can be determined for any input voltage v_s and v_G by moving along the load line between P and S.

(5) For linear amplification, R_L and V_{PP} should be chosen to place the load line in, the region of the plate characteristics where the curves are parallel and equally spaced. The bias voltage, V_{GG}, should be chosen to locate the Q-point in such a way that the peak input voltage of v_s will not drive the grid positive.

(6) The peak-to-peak value of output voltage, v_o', disregarding the dc level, is 220 - 100 = 120 V. This corresponds to a peak-to-peak input of 8 V. A_v = voltage gain of amplifier

$$= \frac{\text{Change in output, volts}}{\text{Change in input, volts}} = \frac{100 - 220}{4 - (-4)}$$

$$= \frac{-120}{8} = -15.$$

Magnitude of ac gain of amplifier

$$= \frac{\text{Output voltage, peak-to-peak}}{\text{Input voltage, peak-to-peak}} , \frac{v_{o,p\text{-}p}}{V_{s,\,p\text{-}p}}$$

$$\frac{120}{8} = 15.$$

This is a dimensionless number since it is the ratio of volts to volts. Note: The voltage gain A_v is negative, because an increase in v_s caused a decrease in v_o', and it is obvious from the waveforms of v_o' and v_s that phase inversion occurs. That is, there is a 180° phase shift between input and output. This is the meaning of the negative sign in $A_v = -15$.

A-9. MAXIMUM OUTPUT VOLTAGE AVAILABLE FROM AN AMPLIFIER

Consider the equation v_o' $V_{PP} - v_L$ and $v_L = i_P R_L$.

Say v_L can be made 0 V at one extreme of input voltage ($i_P = 0$): then $v_o' = V_{PP} - 0 = V_{PP}$.

Say v_L can be made equal to V_{PP} (large i_P) at other extreme: then $v_o' = V_{PP} - V_{PP}$ 0.

Then V_o' peak-to-peak $= V_{PP} - 0 = V_{PP}$.

Thus to increase the output voltage, as large a value as possible should be used for V_{PP}, within the tube voltage capabilities.

A-10. USE OF A BLOCKING CAPACITOR

The output voltage v_o' was "shifted up" by a dc value of 160 V (see waveform for v_o' in Fig. A-9(e)). In an ac amplifier this dc level is objectionable, especially if the output of one tube is to be applied to the grid of another tube. Hence a blocking capacitor is used to keep this dc level from appearing at the output as shown in Fig. A-10.

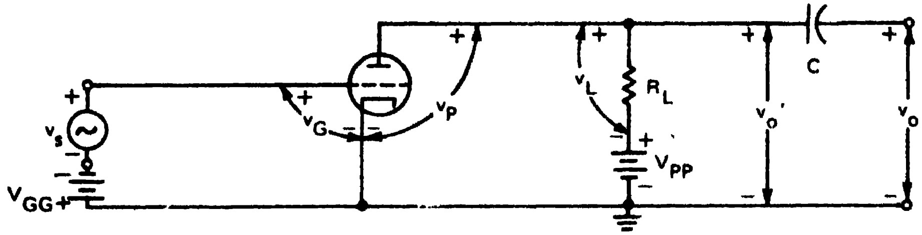


Fig. A-10. Ac amplifier showing output blocking capacitor.

This capacitor is also known as a coupling capacitor, and its value will depend upon the load coupled to the amplifier as well as the expected frequency of operation, as covered in Appendix C. Note that v_L, v_o' and v_o have the same peak-to-peak value, but it is preferable to use the voltage output with respect to ground in order to have a common ground connection (for convenience) between input and output.

A-11. EQUIVALENT CIRCUIT FOR A SMALL SIGNAL AC AMPLIFIER, GROUNDED CATHODE

If the characteristic curves for a tube are available, the gain may be calculated using a load line analysis. However, if only the μ and r_p of the tube are known, the voltage gain can be calculated using an equivalent circuit of the Thevenin type shown in Fig. A-11.

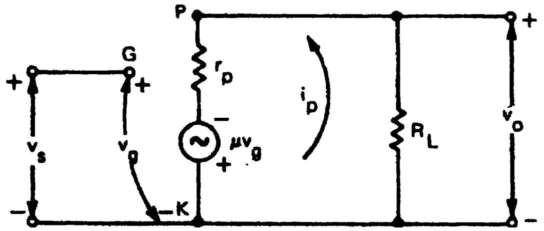


Fig. A-11. Thevenin equivalent circuit for a small signal ac amplifier.

Since $\mu = \Delta v_P/\Delta v_G$, then $\Delta v_P = \mu\Delta v_G$. That is, a change in grid voltage is reflected as a change in the plate circuit of $\mu\Delta v_G$ volts of opposite polarity, shown by the inverted minus and plus signs on the ac voltage generator. This is in series with the plate resistance of the tube r_p. It is required to find an equation for voltage gain, $A_v = v_o/v_g = v_o/v_s$. Note that this equivalent circuit does not include the bias voltage or the V_{PP} supply; it is good only for calculating the gain of a small signal ac amplifier.

$$v_o = -i_p R_L \text{ } (i_p \text{ is the ac component of } i_P). \tag{A-6}$$

$$\text{But } i_p = \frac{\mu v_g}{r_p + R_L}. \tag{A-7}$$

$$\therefore v_o = -i_p R_L = -\frac{\mu v_g}{r_p + R_L} \times R_L$$

$$= -\frac{\mu R_L}{r_p + R_L} \times v_g. \tag{A-8}$$

$$\therefore A_v = \frac{v_o}{v_s} = \frac{v_o}{v_g} = -\frac{\mu R_L}{r_p + R_L} \times \frac{v_g}{v_g}$$

$$= -\frac{\mu R_L}{r_p + R_L}. \tag{A-9}$$

If the load is not a pure resistance R_L but is Z_L, then,

$$A_v = -\frac{\mu Z_L}{r_p + Z_L}. \tag{A-10}$$

An alternate form for A_v is:

$$\frac{\mu R_L}{r_p + R_L} : \frac{-\mu}{\frac{r_p}{R_L} + \frac{R_L}{R_L}} \quad \frac{-\mu}{1 + \frac{r_p}{R_L}} \tag{A-11}$$

The negative sign indicates phase inversion between input and output.

This very important equation shows how the *gain* of an amplifier depends upon the *amplification* factor, mu, of the tube and the plate load resistor. For example, say R_L is numerically equal to r_p:

$A_v = -\mu/(1 + 1) = -\mu/2$. Then the gain of the amplifier is one-half of mu. If R_L 10 r_p,

$$\frac{-\mu}{1 + 10r_n} \quad \frac{-\mu}{1 + 0.1} \quad \frac{-\mu}{1.1} = -0.9\mu.$$

Thus as R_L increases, so does the gain, approaching μ as a maximum. Hence a voltage amplifier should have a large load resistor, as shown by the graph in Fig. A-12. However, since the power delivered to the load decreases with increasing R_L, the load resistor is usually made 2 – 3 times the size of r_p.

The load resistor referred to above is the plate load resistor, R_L, and should not be confused with the load connected to the amplifier, R_o, which was referred to in Chapter 7. The effects of adding a load to the amplifier are considered in the next appendix.

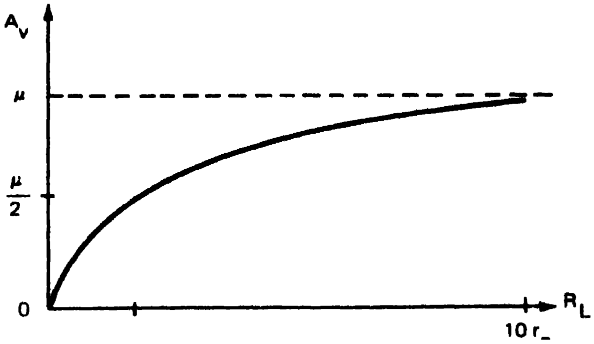


Fig. A-12. Graph showing the variation in voltage gain with load resistance.

A-12 SUMMARY

(1) A triode is basically a modified diode with a grid added to provide control of plate current by means of a voltage on the control grid.

(2) A triode's characteristics consists of a series of curves relating plate current to plate voltage as a function of grid voltage.

(3) The three basic properties of a tube are described by the amplification factor μ, the plate resistance r_p, and the mutual conductance g_m in the following manner:

$$= \frac{\Delta V_P}{\Delta V_G}, \; I_P \text{ constant}; \; r_p = \frac{\Delta V_P}{\Delta I_P},$$

$$V_G \text{ constant}; \; g_m = \frac{\Delta I_P}{\Delta V_G}, \; V_P \text{ constant};$$

$$\mu = r_p g_m.$$

(4) A tube's permissible area of operation on its characteristic curve is bounded by the maximum plate voltage and the maximum plate dissipation hyperbola.

(5) The amplifying action of a triode relies upon the negative grid voltage controlling the amount of plate current through the

load resistor. This current is supplied by the dc plate voltage, V_{PP}.

(6) The analysis of a dc and unloaded ac amplifier may be done by drawing a dc load line on the tube's characteristic curves. One end of the line is at: ($V_P = V_{PP}$, $I_P = 0$) and the other at: ($V_P = 0$, $I_P = V_{PP}/R_L$).

(7) The gain of a simple dc amplifier is the ratio of the change in the output voltage to the change in the input voltage.

(8) In order to prevent the grid from becoming positive in a true class A ac amplifier, a negative grid bias voltage is necessary which is a dc voltage equal to or greater than the peak of the ac input signal.

(9) If the grid is driven positive, we term the operation class A_2, which has the disadvantage of requiring the signal source to provide current with a resulting decrease in input voltage due to the source's internal resistance.

(10) The maximum peak-to-peak output voltage available from an amplifier is limited by the dc supply voltage.

(11) A blocking or coupling capacitor at the output of an ac amplifier serves to couple only the ac component to the load.

(12) A Thevenin equivalent circuit for a small signal ac grounded cathode amplifier provides a means of obtaining the voltage gain using the following equation:

$$A_v = -\frac{\mu R_L}{r_p + R_L},$$

where the negative sign indicates phase inversion.

(13) A voltage amplifier utilizing a triode should employ as high a value of load resistance as possible (to provide a gain approaching μ), but the load is usually made 2-3 times the size of r_p for practical purposes.

QUESTIONS

QA-1. What characteristic does a triode have when operated with its grid at 0 V with respect to the cathode?

QA-2. Explain in your own words what you understand by the term *amplification factor*, as applied to a tube.

QA-3. What does a set of transfer curves consist of?

QA-4. Explain how the mutual conductance may be obtained very easily from the transfer curves and why it is thus called transconductance.

QA-5. If two tubes have the same value of amplification factor, but tube A has a lower value of plate resistance than tube B, what does this mean as far as the comparative values of the tubes' transconductance?

QA-6. What is a maximum dissipation hyperbola?

QA-7. When a triode is used as a dc amplifier, an increased negative voltage on the grid (from the input) *reduces* the flow of current in the plate circuit. Explain how this can cause an increase in output voltage and thus provide voltage gain.

QA-8. What two extremes are used to plot the dc load line on the characteristic curves of an amplifier's active device?

QA-9. What are some of the advantages of not driving the grid of a tube positive or into cutoff? What class of operation is this called?

QA-10. What is the purpose of a grid bias voltage? Is it positive or negative? What must be the minimum value of this bias voltage when amplifying a sinusoidal signal having a peak-to-peak voltage = V?

QA-11. What does it mean to say that an amplifier has a voltage gain $A_V = -20$? Explain both the 20 and the negative sign as it applies to an ac amplifier.

QA-12. What is the largest peak-to-peak output voltage you could hope to get from a class A amplifier operating with $'_{PP}$ 300 V?

QA-13. What is the purpose of a blocking capacitor? What is it also known as?

QA-14. With reference to a Thevenin equivalent circuit for a tube amplifier and the resulting equation for voltage gain, differentiate clearly between voltage gain and amplification factor.

QA-15. Under what conditions does an amplifier have a voltage gain close in value to the tube's amplification factor?

PROBLEMS

PA-1. The characteristics of a tube are to be determined with the aid of the circuit in Fig. A-2. It is found that a plate voltage of 215 V causes a plate current of 10 mA when the grid voltage is -6 V. In order to keep the plate current constant at 10 mA it is necessary to increase the plate voltage to 252 V when the grid is made -8V. It is then found that a further increase in the negative grid voltage to -10 V causes a reduction in plate current to 5.2 mA if the plate voltage is held constant at 252 V.

(a) Determine the amplification factor.
(b) Determine the transconductance.
(c) Obtain the plate resistance.

PA-2. Using the curves in Fig. A-3(a), obtain the 6CG7's characteristics in the vicinity of 150 V plate voltage and a grid voltage of -2 V. Compare your values of μ, r_p and g_m with those in Fig. A-3(b) in the same region.

PA-3. Both sections of a 6CG7 are to be operated near full power capability. Since the total power dissipation capability is 5.7 W, draw the maximum dissipation hyperbola for 2.8 W on the characteristic curves in Fig. A-6. Also indicate the maximum voltage capability of 330 V.

PA-4. A 6CG7 is to be used in a dc amplifier circuit like that of Fig. A-5. It is required to amplify the change that takes place in V_s from 6 V to 2 V. A supply voltage of 320 V is decided upon, but a question remains as to what value to use for R_L.

(a) Determine the voltage gain of the amplifier using the characteristic curves of Fig. A-6 by drawing load lines for the following three values for R_L:

(i) 12 K ohms, (ii) 16 K ohms,
(iii) 20 K ohms.

(b) What conclusion can you draw regarding the selection of R_L for maximum voltage gain?
(c) What was the tube dissipation for R_L = 12 K ohms, V_s = 2 V? How does this compare with the dissipation for R_L = 20 K ohms, V_s = 2 V?

PA-5. Using the maximum dissipation hyperbola drawn in PA-3, determine:

(a) The minimum value of R_L that could be tolerated in the circuit of Fig. A-5 with V_{PP} = 320 V and the tube a 6CG7. (You will have to extend the vertical axis in Fig. A-6.)
(b) The grid voltage that would cause maximum dissipation in the tube.

PA-6. The circuit of Fig. A-7 is to be used as an ac amplifier to provide gain for a 4 V peak-to-peak sinusoidal signal. $V_{PP} = 320$ V and $R_L = 16$ K ohms. The tube is a 6CG7.

(a) What must be the minimum value of bias voltage, V_{GG}?
(b) Calculate the voltage gain when $V_{GG} = -2$ V, -4 V, -6 V, and -12 V.
(c) Sketch the waveform you would expect the output to have, (v_o'), when $V_{GG} = -12$ V.

PA-7. Repeat PA-6 with R_L changed to 100 K ohms, and compare the voltage gains and the waveforms in part (c). What can you conclude regarding the choice of R_L in an ac voltage amplifier as it affects voltage gain and waveform distortion?

PA-8. The circuit of Fig. A-10 is to be used with a 6CG7 to provide a minimum output of 300 V peak-to-peak, with little concern for distortion. Select values for V_{PP}, R_L, V_{GG}, and v_s to provide this output, stating the effective voltage gain. Compare the amplitude of the positive peak with that of the negative peak of the output.

PA-9. The triode section of a 6AV6 tube has an amplification factor of 100 and a plate resistance of 80 K ohms when operated at a plate voltage of 100 V. Determine the voltage gain when this tube is used in an amplifier circuit like that of Fig. A-10, with

(a) $R_L = 100$ K ohms.
(b) $R_L = 80$ K ohms.
(c) $R_L = 470$ K ohms.

PA-10. A 6CG7 has an amplification factor of 20 and a plate resistance of 6.7 K ohms when operated at a plate voltage of 90 V. It is used in an amplifier circuit like that in Fig. A-10. If $R_L = 8$ K ohms, and $C = 0.02\,\mu$F, determine:

(a) the small signal open circuit voltage gain.
(b) the voltage across C if the output is accidentally shorted to ground when a 1 V peak-to-peak signal is applied at 1 kHz.
(c) What will be the output in (b) if the frequency is 10 kHz?

PA-11. If a tube with an amplification factor of 100 and a plate resistance of 60 K ohms is to be used in an amplifier to provide a voltage gain of -80 for small signals, determine the required value of load resistance. R_L.

APPENDIX

B

ANALYSIS OF CLASS A TRIODE AMPLIFIERS

B-0. INTRODUCTION

The separate bias supply used in the previous ac amplifier will work, but a more economical method of self-biasing is possible and is used in most class A amplifiers. The self biasing resistor requires a bias line to be drawn on the characteristic curves to determine the operating point through which the ac or dynamic load line is drawn. Using this line, the harmonic distortion can be calculated and the output power determined. Two other amplifier configurations are also examined that have counterparts in transistor circuitry.

B-1. EVOLUTION OF SELF BIAS

It was shown that it was necessary to make the grid negative by some value of dc voltage with respect to the cathode to place the operation of the tube in the center of the characteristic curves. A separate supply in series with v_s was used for this, as shown in Fig. B-1(a).

In (b) the battery has been moved into the cathode line, making the cathode positive with respect to the grid. This is exactly the same as in (a) from a bias standpoint, although V_{GG} must now be capable of handling the plate current of the tube, i_P. Also, it is assumed the internal resistance of V_{GG} is negligible, and a dc path exists through v_s.

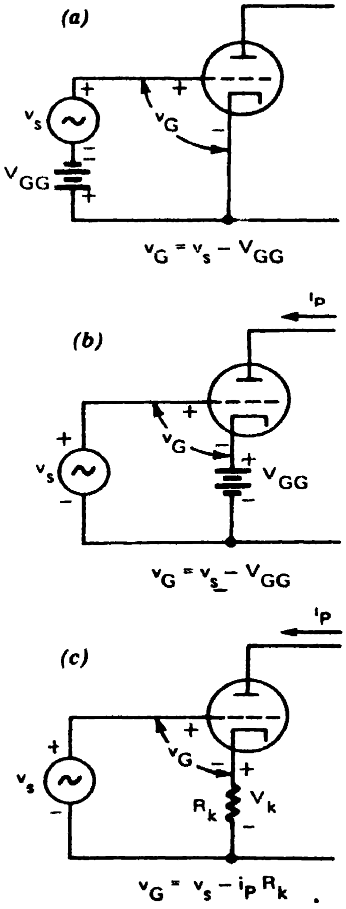


Fig. B-1. Diagrams showing the evolution of self bias.

Since the circuit in (b) makes the cathode positive with respect to ground, the same effect would take place if the dc plate current I_P were made to flow through a resistor placed in the cathode line, as shown in (c). Then

$$V_k = V_{GG} = I_P R_k . \qquad \text{(B-1)}$$

This would give the correct dc bias on the grid when $v_s = 0$. That is, $V_G = -I_P R_k$. However, what is the effect of this self bias when an ac signal, v_s is applied?

Since $v_G = v_s - i_P R_k$ (note that small i_P, not I_P, is used since now the plate current will vary with time and is not constant), the voltage from grid to cathode, v_G, will vary due to two sources. Obviously, v_s is the input signal and is applied to make i_p vary and develop an output in R_L. However, when i_p increases due to an increase in v_s the product $i_p R_k$ will also increase, tending to offset the increase in v_s. Hence, v_G will *tend* to remain constant. It will increase a small amount, and some output voltage will result but will be much less than with separate bias. This is negative or degenerative feedback and always acts so as to reduce the output.

B-2. BYPASS CAPACITOR

How can this drawback be eliminated? Consider the plate current when an ac signal is applied. This is shown pictorially in Fig. B-2. The plate current should be thought of as a dc component, I_P, and an ac component, i_p. Since a constant dc bias is required, all that is necessary is to prevent the ac component from flowing through R_k, allowing only the dc component I_P to develop bias. This is done using a cathode bypass capacitor, C_k, as illustrated in Fig. B-3. The reactance of C_k is made small compared with R_k at the lowest frequency of operation to provide adequate bypass for all frequencies.

$$\frac{R_k}{10} \qquad 2\pi fC_L \tag{B-2}$$

The capacitor is fairly large, ranging from 5 μF to 50 μF, depending upon R_k. Accordingly, electrolytic capacitors are normally used and must be placed in the circuit with correct polarity. Nevertheless, at low frequencies the reactance of C_k may become too high to give good bypass of the ac component around R_k. Hence some ac voltage may de developed across R_k and cause a reduction in voltage gain due to negative feedback. However, at

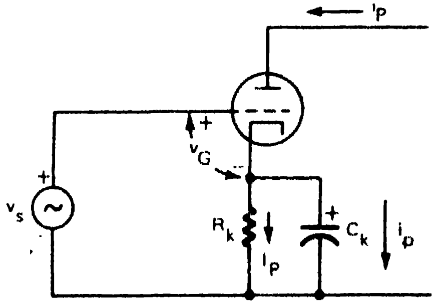


Fig. B-3. Showing the bias resistor and bypass capacitor.

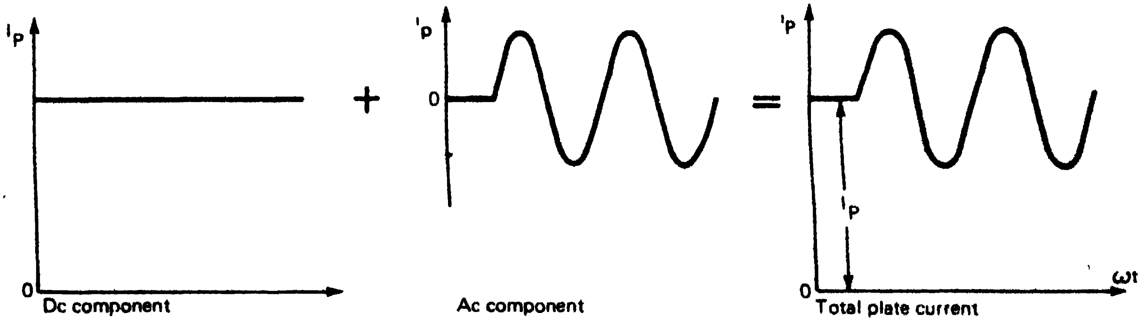


Fig. B-2. Showing how the total plate current consists of two components.

the low voltage involved, it is quite easy to get a large enough capacitance in a small enough physical size to avoid appreciable reduction in gain at low frequencies due to this factor.

B-3. GRID-LEAK RESISTOR

Whether separate bias or self bias is used, it has been assumed that the signal source v_s has a dc path by which the negative bias can be applied to the grid. This is not always the case or, if it is available, what would be the effect of removing the source? Consider Fig. B-4.

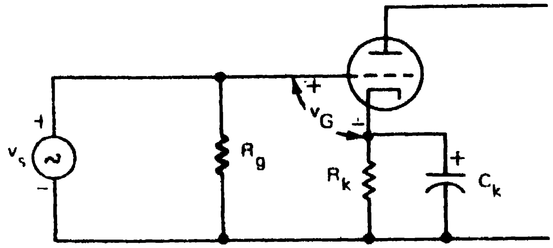


Fig. B-4. Showing the grid-leak resistor, R_g.

If v_s is removed (without R_g), the grid becomes a "floating grid"; normally the grid voltage would become zero, and the plate current would perhaps increase to a damaging level. The addition of resistor R_g provides a path by which the bias voltage developed across R_k is always connected to the grid and holds it negative. However, R_g is made large, 470 K ohms to 2 megohms, so that the grid is still effectively isolated from ground. R_g is also called a grid-leak resistor, since it provides a path for electrons to leak off the grid to ground. Otherwise, it is possible (without R_g) for the grid to accumulate sufficient electrons (on their way to the plate) so that the tube tends to become cut off, and operation becomes erratic.

Note that no dc current flows in R_g, just a small amount of ac current due to v_g. However, there is a method of biasing, called grid-leak bias, used in class C RF amplifiers, which uses the applied signal to develop bias. This requires the addition of a series capacitor and necessitates driving the grid positive to draw grid current.

B-4. DETERMINATION OF OPERATING POINT

The addition of self bias raises the question as to where is the operating point located? This is not directly answerable as can be seen from the following paradox.

If the amount of bias developed is dependent upon the amount of plate current and the plate current in turn depends upon the bias, where does one start? The answer is to draw the dc load line and then draw a *bias line*, the intersection of the two giving the Q point.

The first step is to draw the dc load line. As before, one end of the dc load line is located on the characteristic curves at $i_P = 0$, $v_P = V_{PP}$ (cutoff). The other is located at

$$v_P = 0, \; i_P = \frac{V_{PP}}{R_L + R_k} \quad \text{(saturation).}$$

This is shown in Fig. B-5 for a numerical example, in which $R_L = 20$ K ohms, $R_k = 1$ K ohm and $V_{PP} = 300$ V.

Note that R_k is added to R_L to obtain the resistance value for the load line, since the dc current cannot pass through C_k.

B-5. BIAS LINE

To determine the bias line, assume a plate current, say $i_P = 1$ mA. This will develop a bias $V_G = -i_P R_k = -1 \times 10^{-3} \times 1 \times 10^3 = -1$ V. Drawing a line horizontally on the plate characteristics at $i_P = 1$ mA, estimate where the bias line v_G = -1 V is located and indicate a point here. This is nowhere near the dc load line, so assume another current, say i_P = 3 mA; this gives v_G = -3 V. Plot this point. Try i_P = 5 mA; this gives v_G = -5 V. Draw a line through these points. This will not be exactly a straight line; however, if two plotted points are used, one on either side of the dc load line and close to it, a straight line drawn between the two will give a sufficiently accurate intersection and will determine the Q point. This operating point is indicated by giving I_P and V_P and V_G. See Fig. B-5(b).

In this case, the bias voltages were equal to the plate current (numerically) only because R_k was 1 K ohm. This will not always be true. In this case V_G turned out to be exactly -4 V. If the intersection falls between two given curves the value of bias will have to be estimated.

It can be argued that the intersection of this bias line and dc load line *is* the operating point by assuming the operating point to be first higher, then lower, than the indicated Q point. In each case the plate current is either too high to develop the necessary bias, or too low. Hence the indicated Q point is the only stable place at which conditions balance to answer the paradox mentioned earlier.

B-6. AC LOAD LINE

For the circuit in Fig. B-5(a), what will be the effect on the dc load line of applying an ac signal, v_s? Say an 8 V peak-to-peak signal is applied. With a bias of -4 V, this will drive the grid up to 0 V and down to -8 V, but this will *not* be along the dc load line (as it was for separate bias). A new load line must be drawn, since for ac the capacitor C_k bypasses R_k, so that R_L is the only load seen by the ac signal. This ac load line (or dynamic load line since this is "where the action is") must pass through the operating point found earlier and, by a previous analysis (Section 2-1), must have a slope of $\frac{-1}{R_T}$. The question then is, what is meant by drawing a line having a slope = $\frac{-1}{20\text{ K ohms}}$?

The negative sign indicates the line must slope from the upper left hand corner of the characteristics to the lower right hand corner. A positive slope would be upward from the left to right. The larger the number, the greater the steepness, as illustrated by Fig. B-6.

Slope of ac load line

$\frac{-1}{20\text{ K ohms}}$. This is $\frac{1}{\text{kilohms}}$ - 10^3 ohms

$\frac{1 \times 10^{-3}}{\text{ohms}}$ Since V = IR, then:

$\frac{1}{R} = \frac{I}{V}\ \frac{\text{amps}}{\text{volts}} \therefore \frac{1}{20\text{ K}} - 1 \times \frac{10^{-3}}{20}$ has units of

$\frac{1\text{ mA}}{20\text{ V}}$.

Then the ac load line slopes downward from left to right having a slope of 1 mA per 20 V. Note:

$$\text{Slope} = \frac{\text{rise}}{\text{run}} = -\frac{\text{vertical}}{\text{horizontal}}\ \ \frac{\Delta i_P}{\Delta v_p}$$

This means that with respect to any reference point on the line, a horizontal movement of 20 V followed by a vertical movement of 1 mA will come to rest on the line at a different point. Note that a slope of 1 mA per 20 V is the same as 3 mA per 60 V or 5 mA per 100 V. Generally a slope of 5 mA per 100 V is easier and more accurate to draw than is 1 mA per 20 V.

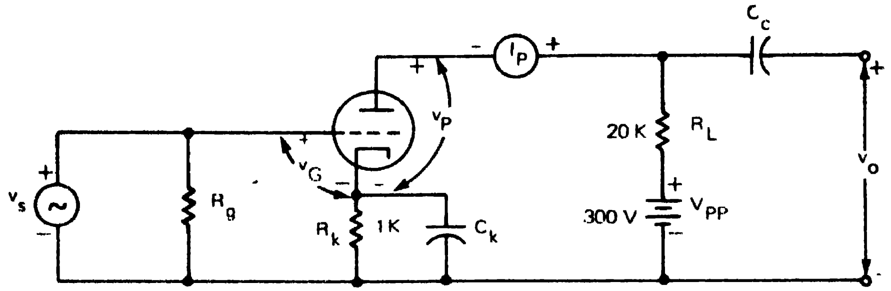


Fig. B-5. (a) Circuit of an ac amplifier to be analyzed.

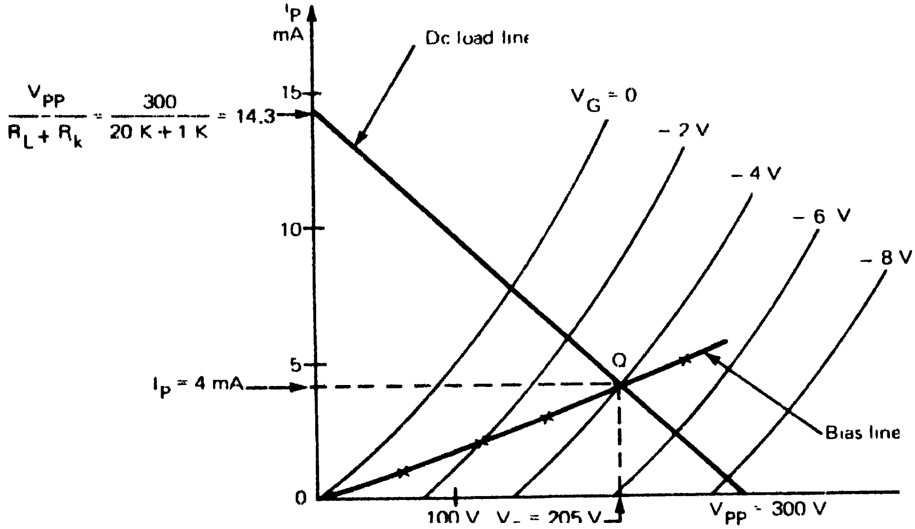


Fig. B-5. (b) Showing the dc load line and bias line for an ac amplifier.

Now consider drawing the ac load line on the characteristic curves, using the Q point as the reference point, since the determination of one other point will allow the drawing of the load line. See Fig. B-7.

Starting at Q, V_P = 205 V (for example) and I_P = 4 mA. Move to the *left* 100 V horizontally; using the scale for v_P, to v_P = 105 V. This is Δv_P. Now move vertically 5 mA using the scale for i_P, to i_P = 4 mA + 5 mA = 9 mA. This is Δi_P. This establishes another point for the ac load line, S. Now draw a straight line through S and Q yielding the ac load line. In this case, the ac or dynamic load line is not very different from the dc load line since R_k = 1 K ohm did not make much difference compared with R_L = 20 K ohms. However, it can be seen that driving the grid between 0 V and -8 V will produce slightly less output voltage than the dc load line would give. Note that the same line would result if v_P was taken 100 V to the *right* to v_P = 305 V and Δi_P *down* 5 mA to -1 mA.

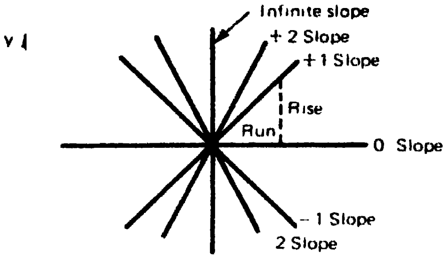


Fig. B-6. Sketch showing the meaning of different slopes.

B-7. EFFECT OF AN AC LOAD

Normally, the amplifier is used to drive some load, R_o, as shown in Fig. B-8.

For a numerical example, let $R_o = 20$ K ohms with the previous values for R_k, R_L and V_{PP}.

The dc load line and bias line will be the same as before to give the same operating point. But the ac load line will be different, since at the frequency under consideration it is assumed that the reactance of C_c is negligible, making R_o in parallel with R_L as far as ac signals are concerned. The effective ac load resistance is:

$$R_e = R_L \| R_o = \frac{R_L R_o}{R_L + R_o} \tag{B-3}$$

For $R_L = R_o = 20$ K ohms, this represents an effective load of 10 K ohms. The slope of the ac load line is now

$$\frac{-1}{R_e} = \frac{-1}{10 \text{ K ohms}}.$$

$$\text{Slope} = \frac{1}{10 \text{ K ohms}} = \frac{1 \text{ mA}}{10 \text{ V}} = \frac{10 \text{ mA}}{100 \text{ V}} = \frac{\Delta i_P}{\Delta v_P}$$

Moving to the left 100 V from Q and vertically 10 mA gives point S and hence the ac load line, shown in Fig. B-9. Note that it is twice as steep as the dc load line and V_o peak-to-peak is less with R_o connected than when there is no load (the earlier ac load line in Fig. B-7 is used for no load). The larger the load connected (the smaller R_o is) the steeper will be the ac load line, with a resulting decrease in output voltage and gain.

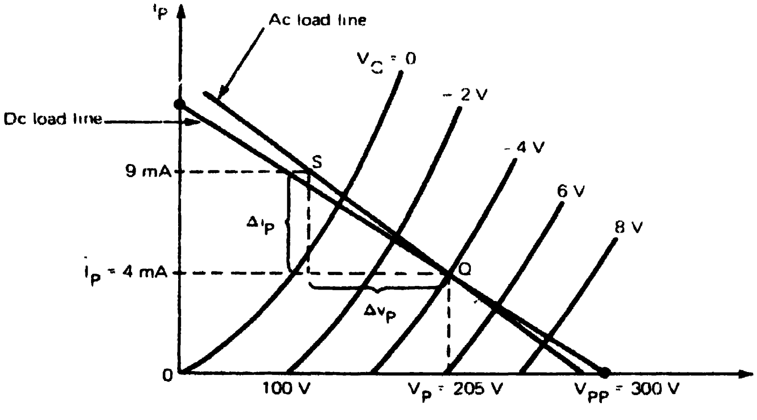


Fig. B-7. · Showing how to locate point S for drawing the ac load line.

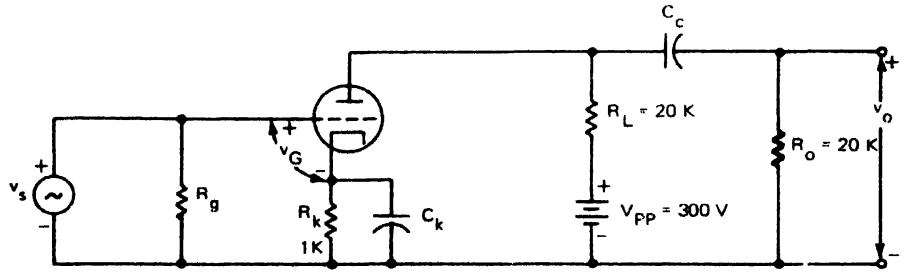


Fig. B-8. Ac amplifier with load, R_O, coupled to output.

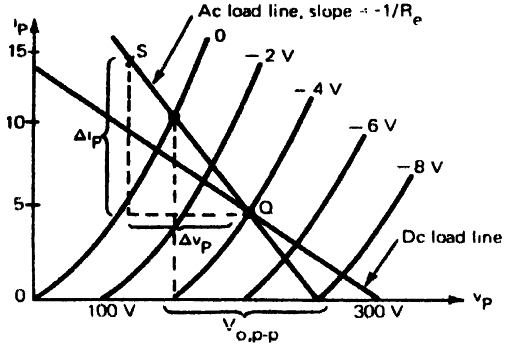


Fig. B-9. Ac load line drawn for an effective load resistance, R_e.

(Note that the peak value of current through the tube also increases.)

The decrease in gain that results when a load is added to an amplifier may be seen by referring to the gain equation for the small signal equivalent circuit, developed in Appendix A-11.

$$A_v = \frac{\mu R_L}{r_p + R_L} \tag{B-4}$$

However, this is the unloaded gain of the amplifier. If a load R_o is connected and the amplifier operated at a frequency such that the reactance of C_c is negligible compared with R_o, then R_L and R_o are effectively in parallel as far as ac is concerned. This combination has been referred to as R_e. Thus:

$$A_v = \frac{\mu R_e}{r_p + R_e} \tag{B-5}$$

Example B-1. As a numerical example, consider the triode in Fig. B-8 to have a μ = 21 and r_p = 6 K ohms. Determine the open circuit voltage gain with R_L = 20 K ohms, and the voltage gain when R_o = 20 K ohms is added to the output. Assume reactance of C_c is negligible.

Solution

At no load: $A_v = \frac{\mu R_L}{r_p + R_L} = \frac{21 \times 20\text{ K}}{6\text{ K} + 20\text{ K}} = 16.2.$

With load: $A_v = \frac{\mu R_e}{r_p + R_e} = \frac{21 \times 10\text{ K}}{6\text{ K} + 10\text{ K}} = 13.1.$

Note that an effective load change of one-half, from 20 K ohms to 10 K ohms did not reduce the voltage gain by one-half.

B-8. EFFECT OF COUPLING CAPACITOR C_c

When no load was connected to the amplifier, the value of C_c was unimportant. It was just used to block the dc voltage from whatever measuring instrument was connected across the amplifier. However, when an ac voltage is developed across R_L (ahead of the capacitor), the reactance of C_c will determine how much voltage will appear across R_o.

If, at a lower frequency, X_{C_c} becomes comparable to R_o, then v_o will drop in value below what the ac load line will predict, since this is drawn on the basis of negligible reactance. The coupling capacitor reactance is one of the main sources for loss of voltage gain in RC-coupled amplifiers at low frequencies and will be covered in detail in the next chapter.

B-9. PERFORMANCE CALCULATIONS FOR A CLASS A TRIODE AMPLIFIER

Let us examine the conversion efficiency, harmonic distortion, and power gain by means of a numerical example.

Example B-2 For the amplifier in Fig. B-10(a) and using the characteristic curves in Fig. B-10(b), determine:

(a) The value of R_k that will allow the amplifier to work in true class A when an input signal of 8 V p-p is applied.
(b) The open circuit voltage gain (no R_o).
(c) The loaded voltage gain (with R_o).
(d) The harmonic distortion.
(e) The conversion efficiency.
(f) The plate dissipation.
(g) The power gain in decibels.

Assume reactances of C_c and C_k are negligible.

Solution. Note that in Fig. B-10(a) the amplifier has been drawn slightly different, with R_L drawn vertically upward to V_{PP}. It is understood that the negative end of V_{PP} is connected to the common line (ground). This is a space saving device for drawing cascaded amplifiers all working from the same source of V_{PP}. However, R_L and R_o are still considered to be in paralled for ac.

Bias Resistor

The first step is to draw the dc load line. Since R_k is not yet known, we assume it is small compared with R_L. Thus:

(1) At $i_P = 0$, $v_P = V_{PP} = 300$ V.

(2) At $v_P = 0$, $i_P = \dfrac{V_{PP}}{R_L} = \dfrac{300}{20\text{ K}} = 15$ mA.

The minimum bias voltage required is -4 V since the peak of the input signal is 4 V. (Note: We could use −6 V; see assigned student problem.)

Thus the Q point is: $I_P = 7.2$ mA, $V_P = 155$ V.

We need a resistor to develop 4 V across it due to I_P.

$$R_k = \frac{V_{GG}}{I_P} = \frac{4}{7.2 \times 10^{-3}} = 555 \text{ ohms.}$$

(Use 560 ohms.)

Actually, point (2) of the dc load line is located at:

$$i_P = \frac{V_{PP}}{R_k + R_L} \quad \frac{300}{(0.56 + 20)\text{K}} = 14.6 \text{ mA}$$

and $v_P = 0$, but redrawing the load line changes the Q point a negligible amount.

Open-Circuit Voltage Gain

With no load R_o connected, operation as an ac amplifier will be along the dc load line, from Q to P′, back to Q and down to R′, resulting in an output voltage of:

$V_{o\,,p\text{-}p}'$ 212 - 92 = 120 V peak-to-peak.

This resulted from an input voltage of 8 V p-p.

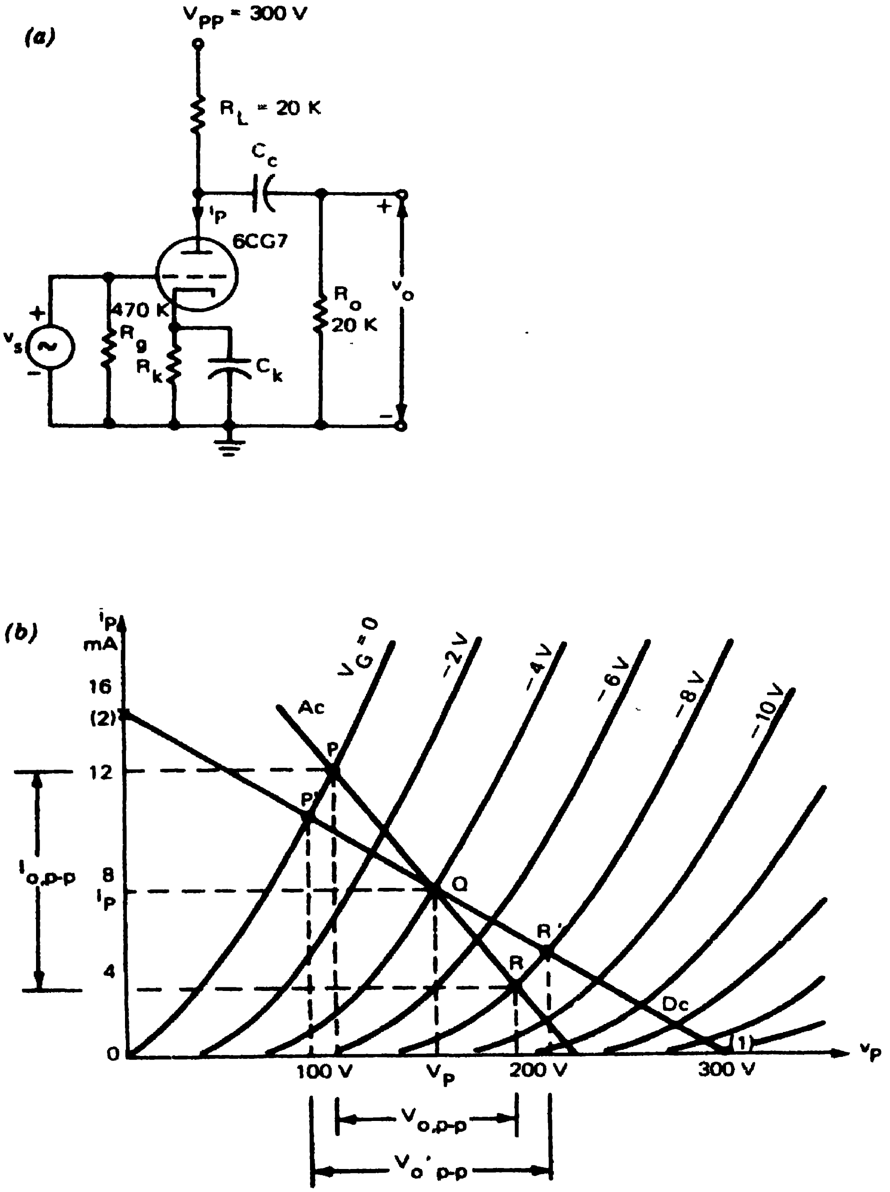

(a)
V_{PP} = 300 V
R_L = 20 K
C_c
i_P
6CG7
+
v_s
−
470 K
R_g
R_k
C_k
R_o
20 K
+
v_o
−
(b)
i_P
mA
16
(2)
12
$I_{o,p\text{-}p}$
8
I_P
4
0
$V_G = 0$
−2 V
−4 V
−6 V
−8 V
−10 V
Ac
P
P'
Q
R
R'
Dc
(1)
100 V
V_P
200 V
300 V
v_P
$V_{o,p\text{-}p}$
$V_o{}'_{p\text{-}p}$

Fig. B-10. (a) Circuit of triode amplifier to be analyzed. (b) Characteristic curves with dc and ac load lines.

$$\text{Thus } A_v' = \frac{V'_{o,p\text{-}p}}{V_{s,p\text{-}p}} = \frac{120}{8} = 15.$$

(Phase inversion is understood to occur in a grounded cathode amplifier, so the minus sign has not been included.)

Loaded Voltage Gain

Adding a 20 K ohm load for R_o requires the ac load line to be drawn with a slope

$$\frac{-1}{R_e} = \frac{-1}{10\text{ K ohms}} \text{ or } 1\text{ mA}/10\text{ V}.$$

Drawing this through Q and letting v_s drive the grid between P and R results in the following values:

$v_G = 0\text{ V}, I_{max} = 12.4\text{ mA} \quad V_{min} = 105\text{ V}.$

$v_G = -4\text{ V}, I_P = 7.2\text{ mA} \quad V_P = 155\text{ V}.$

$= -8\text{ V}, I_{min} = 3\text{ mA} \quad V_{max} = 195\text{ V}.$

$V_{o,p\text{-}p} = V_{max} - V_{min} = 195 - 105 = 90\text{ V}.$

$$\therefore A_v = \frac{V_{o,p\text{-}p}}{V_{s,p\text{-}p}} = \frac{90}{8} = 11.3.$$

Harmonic Distortion

Let us first check the second harmonic distortion using the three-point method:

$$A_1 = \frac{I_{max} - I_{min}}{2} = \frac{12.4 - 3}{2} = 4.7\text{ mA}.$$

This is the amplitude of the fundamental,

$$= \frac{I_{o,p\text{-}p}}{2}.$$

$$A_o = A_2 \quad I_{max} + I_{min} - 2I_P$$

$$\underline{12.4 + 3 - 2(7.2)} \; 0.25\text{ mA}.$$

This indicates that the dc supply current will increase from 7.2 mA to 7.45 mA when a signal is applied.

The percent second harmonic distortion

$$\frac{0.25}{4.7} \times 100\% = 5.3\%.$$

This fairly large amount of second harmonic indicates that perhaps the higher harmonics are not negligible. Let us use the five-point method of Section 7-13, for which we need $I_{1/2}$ and $I_{-1/2}$. These correspond in our case to the times when the grid is half-way between -4 V and 0 V, and when the grid is half-way between -4 V and -8 V, respectively.

From Fig. B-10(b) we obtain:

$I_{1/2} = 9.8\text{ mA at } v_G = -2\text{ V}.$
$I_{-1/2} = 5\text{ mA at } v_G = -6\text{ V}.$

Using these values with the above values for I and substituting in the appropriate equations of Section 7-13, we obtain the following amplitudes and distortion:

$A_o = 0.3\text{ mA}, A_1 = 4.73\text{ mA}, A_2 = 0.25\text{ mA},$
$A_3 = -0.03\text{ mA}, A_4 = -0.05\text{ mA}.$

$$D_{,} \quad \frac{0.25}{4.73} \times 100\% = 5.3\%,$$

$$D_3 = \frac{0.03}{4.73} \times 100\% = 0.63\%,$$

$$D_4 = \frac{0.05}{4.73} \times 100\% = 1.05\%.$$

$$D = \sqrt{D_2{}^2 + D_3{}^2 + D_4{}^2}$$

$$= \sqrt{5.3^2 + 0.63^2 + 1.05^2} = 5.4\%.$$

Apparently our three-point method was quite accurate in this case in determining our distortion but gave slightly different values for our fundamental and the increase in average current.

Conversion Efficiency

Plate circuit conversion efficiency

$$= \frac{\text{Total ac output power}}{\text{Total dc input power}} = \frac{P_o'}{P_{dc}}.$$

The total ac output power is that developed in R_L and R_o, or in this case, $R_e = 10$ K ohms.

Amplitude of the fundamental $= A_1 = 4.73$ mA (slightly different from $I_{o,p\text{-}p}/2$).

Thus

$$\frac{A_1^2}{2} \times R_e \quad \frac{(4.73 \times 10^{-3})^2}{2} \times 10 \times 10^3$$

$$112 \text{ mW.}$$

P_{dc} is the total power supplied by the V_{PP} supply at an average current I_P:

$$P_{dc} = V_{PP} I_P = 300 \times 7.2 \times 10^{-3} = 2.16 \text{ W.}$$

Thus

$$\eta = \frac{P_o'}{P_{dc}} = \frac{0.112}{2.16} \times 100\% = 5.2\%.$$

It will be shown later that the maximum theoretical efficiency for this type of circuit is 25%, although typical values usually run closer to 15%. The above low value of 5.2% results from a fairly high no-signal dissipation in the tube compared with the output power.

Plate Dissipation

The dc input power to the tube itself is given by $V_P I_P = 155 \times 7.2 \times 10^{-3} = 1.12$ W, and represents the heat dissipated in the tube at quiescent conditions (no signal).

Question: Does the power dissipated by the tube increase or decrease when a signal is applied?

Answer: This must be answered carefully since there are two factors to consider. We have seen how the average current through the tube increases from 7.2 mA to 7.5 mA when the signal is applied. However, the increased dc voltage drop across R_L will reduce the plate voltage from 155 V to 149 V. That is, the operating point "tends" to move up the dc load line. But the increased current through the tube will develop a larger voltage across R_k, the bias resistor, causing the negative grid voltage to change from -4 V to -4.17 V. Consequently, the operating point does not move up the load line significantly; it stabilizes at a point very close to its initial location. Hence no increase in tube dissipation of any significant amount results from this quarter, which incidentally only had to be considered because distortion occurred when the signal was applied.

Actually, the power dissipated in the tube decreases. This is because the active device converts dc power into ac power at the output, P_o'. Thus the dissipation in the tube will decrease by this amount, from 1.12 W to 1.12 - 0.112 W or approximately 1 W. Thus the active device is cooler under dynamic conditions than in static conditions; this is always the case in a class A amplifier.

Power Gain

Although this amplifier is not designed as a power amplifier, we may calculate its decibel power gain.

Since the grid was never driven positive (true class A), the input resistance to the amplifier is determined by R_g, so that $R_i = 470$ K ohms with an input signal v_s of 8 V p-p.

Thus the power supplied by the source is

$$\frac{V_s^2}{R_i} \quad \frac{8^2}{8 \times 470 \times 10^3} \text{ watts} = 17\ \mu\text{W}.$$

The power at the fundamental delivered to the externally connected load, R_o, is P_o and is given by

$$P_o = P_o' \frac{R_L}{R_o + R_L} = 112 \times \frac{20\text{ K}}{20\text{ K} + 20\text{ K}}$$

$$= 112 \times 1/2 = 56\text{ mW}.$$

P_o is one-half of P_o' in this case because $R_o = R_L$, and the ac power divides equally. Note this is the true output power at the fundamental frequency and is different from the value obtained if $V_{o,p\text{-}p}$ is used:

$$\frac{V_{o,p\text{-}p}^2}{8R_o} \quad \frac{90^2}{8 \times 20 \times 10^3} \quad 50.6\text{ mW},$$

since this equation assumes no distortion.

$$\text{Power gain } A_p = \frac{P_o}{P_i} = \frac{56\text{ mW}}{17\ \mu\text{W}} = \frac{56 \times 10^{-3}}{17 \times 10^{-6}}$$

$$= 3.3 \times 10^3.$$

$$\text{dB power gain} = 10 \log A_p = 10 \log 3.3 \times 10^3$$

$$= 35.2\text{ dB}.$$

Note the distinction that has been made between P_o and P_o'. P_o' represents the total ac power in R_o and R_L and was thus used in determining the conversion efficiency. But P_o represents the usable power in the load R_o that is connected to the amplifier and is the value to be compared with P_i.

It is left as a student exercise to show that the alternative equations for calculating decibel power gain using R_i, R_o, V_i, V_o and I_i, I_o yield the same gain. (Note that $V_{o,p\text{-}p}$ at the fundamental is not 90 V; it is $2A_1R_e = 2 \times 4.73 \times 10^{-3} \times 10 \times 10^3 = 94.6$ V p-p.)

B-10. OTHER AMPLIFIER CONFIGURATIONS

In the study of transistor amplifiers there are three common types, each having its own characteristics, depending upon where the output is taken. These are also found in tube amplifiers. The first that we have studied in this chapter is the most common, the grounded-cathode amplifier. It is popular for its good voltage, current, and power gains, and its high input impedance. The others are the cathode follower and the grounded-grid amplifiers.

B-11. CATHODE FOLLOWER

As shown by Fig. B-11(a), the output from a cathode follower is taken across the unbypassed cathode resistor which supplies the

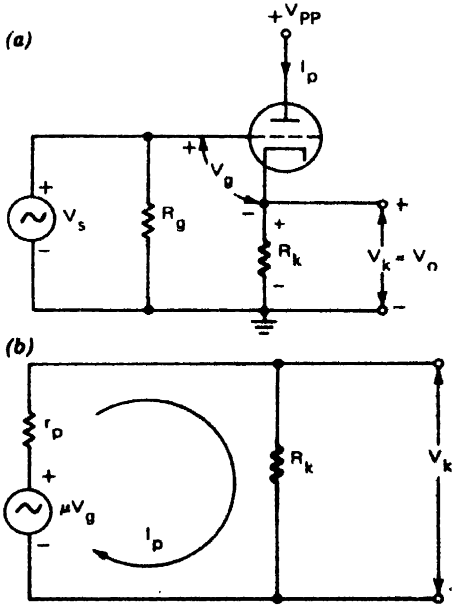


Fig. B-11. (a) Cathode follower amplifier with no external load. (b) Small-signal equivalent circuit for cathode follower.

only load (in addition to r_p) to V_{PP}. Since V_{PP} is essentially a short circuit for ac signals the plate is effectively grounded, so that this amplifier is also known as a grounded-plate amplifier.

Voltage Gain

From Fig. B-11(b) it may be shown that:

$$\frac{\mu R_k}{r_p + (\mu+1)R_k} \qquad \frac{g_m R_k}{1 + g_m R_k}, \qquad \text{(B-6)}$$

if $\mu >> 1$.

Evidently the gain is less than unity and no phase inversion occurs. If $\mu = 19$ and $r_p = 10$ K ohms, $A_v = 0.66$ if $R_k = 1$ K ohm. And $A_v = 0.95$ if $R_k = 10$ K ohms.

Output Resistance

The equivalent circuit of Fig. B-11(b) was drawn with the plate circuit voltage given in terms of the grid voltage V_g. If expressed as a function of the input voltage V_s, it becomes $\mu V_s/(\mu + 1)$, and the resistance r_p becomes $r_p/(\mu + 1)$.

Consequently, the output resistance clearly consists of the following parallel combination:

$$R_{out} \qquad R_k \| \mu + 1 \, . \qquad \text{(B-7)}$$

To use the previous example:

$$R_{out} \qquad 1\text{ K} \| \frac{10\text{ K}}{19 + 1} = 1\text{ K} \| 0.5\text{ K} = 330 \text{ ohms.}$$

Even if $R_k = 10$ K ohms, $R_{out} = 480$ ohms.

Thus, a cathode follower has a high input impedance and a low output impedance, possibly providing impedance matching. It is often used as the last stage of a radio tuner's audio section, in order to be able to drive any amplifier to which it may be connected.

Graphical Analysis

A graphical analysis of the cathode follower in Fig. B-11(a) indicates that the quiescent point is near cutoff, and can accept an input signal of only 30 V p-p. A more practical bias circuit involves breaking R_k into two values of 820 ohms and 9.1 K ohms, and connecting R_g to their junction.* This effectively reduces the bias and allows a more symmetrical operation, so that the maximum input signal that may be accommodated is in the order of 150 V p-p.

B-12. GROUNDED-GRID AMPLIFIER

This amplifier has few applications except as a tuned voltage amplifier at very high and ultrahigh frequencies, where it is often used as the first stage in an FM (frequency modulated) receiver for example, working in the 88 to 108 MHz range. An example of such an amplifier is shown in slightly simplified form in Fig. B-12(a).

R_L actually consists of a tuned resonant LC tank circuit, but is resistive in nature at the frequency of operation. V_s and R_s constitute the input signal from a 300-ohm antenna, which is coupled to the cathode through C_c. Self bias is obtained by R_k but is bypassed by C_k to avoid a reduction in voltage gain. RFC is a radio-frequency choke that presents a high-reactance path to the incoming signal but places the bottom of R_k at ground potential for dc, thus providing the necessary grid-to-cathode bias. The grounding of the grid provides an electrostatic shield between the cathode and plate, and thus prevents coupling between the input and output circuits due to capacitive

---

*See A. Mottershead, *Laboratory Experiments for Electronic Devices and Circuits*, (Pacific Palisades, Calif.: Goodyear Publishing Co., 1973), Experiment B.

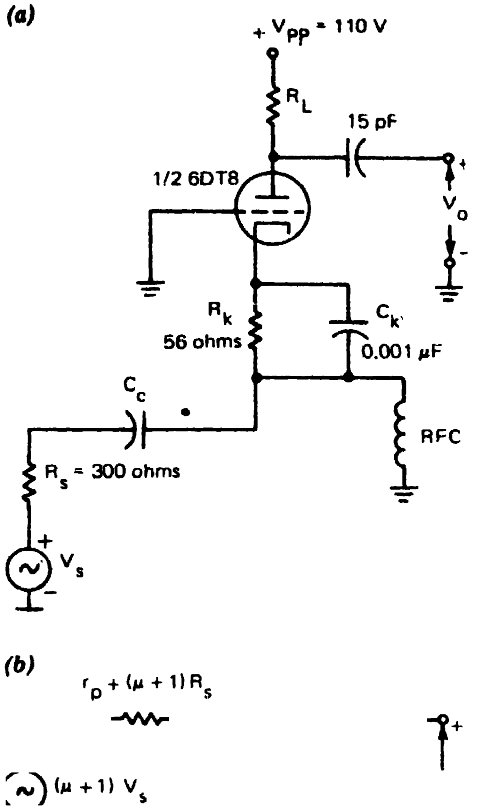


Fig. B-12. (a) Grounded-grid VHF amplifier. (b) Small signal equivalent circuit for grounded-grid amplifier.

effects, allowing operation in the VHF range of 30 to 300 MHz.

It may be shown from the equivalent circuit of Fig. B-12(b) that the voltage gain is given by

$$A_v = \frac{v}{V_s} = \frac{(\mu + 1)R_L}{r_p + R_L + (\mu+1)R_s}. \qquad \text{(B-8)}$$

For a 6DT8 with $\mu = 60$ and $r_p = 15$ K ohms, $R_s = 300$ ohms and R_L assumed to be 10 K ohms, $A_v = 14$.

The input impedance is quite low and is given by

$$R_i \quad \frac{+ R_L}{\mu + 1} \qquad \text{(B-9)}$$

which for the circuit in Fig. B-12(a) amounts to only 410 ohms. However, this matches very closely the antenna's impedance of 300 ohms and so provides a good power transfer.

The output impedance is somewhat higher than the grounded-cathode amplifier and is given by

$$R_o = R_L \Big\| \Big[r_p + (\mu + 1)R_s\Big], \qquad \text{(B-10)}$$

which is 7.7 K ohms for this example.

Finally, no phase inversion occurs between input and output.

B-13. SUMMARY

(1) Self biasing a class A triode amplifier consists of adding a resistor in the cathode line, suitably bypassing it with a capacitor for ac signals and adding a grid-leak resistor from grid to ground.

(2) The operating point of a self-biased amplifier is found from the intersection of the dc load line and the bias line.

(3) When a load is coupled to the output of an amplifier, graphical analysis requires the drawing of an ac load line, slope $= -1/R_e$, through the operating point, or Q point.

(4) The addition of a load to an amplifier reduces the voltage gain from the unloaded value and causes higher peak values of tube current.

(5) Using the ac load line, the harmonic distortion and conversion efficiency may be calculated.

(6) In a class A amplifier, the plate dissipation of a tube is greatest with no signal applied and, when a signal is applied, decreases by an amount equal to the total ac output power from the tube.

(7) A cathode follower has the output taken across the unbypassed cathode resistor and exhibits a voltage gain less than one:

$$A_v = \frac{g_m R_k}{1 + g_m R_k} \quad \text{if } \mu >> 1.$$

(8) A cathode follower's input resistance is very high, but its output resistance is very low and it is thus a good impedance matching device:

$$R_{out} \quad R_L \quad \frac{r_p}{\mu + 1}$$

(9) A cathode follower can accept very large input signals without grid current or cutoff, typically 10 times as large as for a grounded-cathode.

(10) A grounded-grid amplifier, usually used in VHF applications, has a low input resistance, a high output resistance, and a voltage gain that depends upon R_s:

$$\frac{(\mu + 1)R_L}{r_p + R_L + (\mu + 1)R_s}; R_i \quad \frac{r_p + R_L}{\mu + 1}$$

$$\| \left[r_p + (\mu + 1)R_s \right].$$

QUESTIONS

QB-1. Describe in your own words how the three circuits in Fig. B-1 provide the same effective bias voltage.

QB-2. What is the drawback of the circuit in Fig. B-1(c)? How is this overcome?

QB-3. What is the purpose of adding the grid-leak resistor, R_g, in Fig. B-4?

QB-4. How does the addition of a cathode resistor modify the drawing of the dc load line?

QB-5. How is the operating point determined in an amplifier using self bias?

QB-6. If the effective ac load for an amplifier is 20 K ohms explain how you would draw an ac load line through the operating point having a slope equal to -1/20 K ohms.

QB-7. What is the effect on the ac load line of connecting a load to the amplifier? What effect is there on the voltage gain?

QB-8. Explain how the coupling capacitor can cause a reduction in voltage gain at low frequencies?

QB-9. Explain why a triode used in a class A amplifier is cooler under dynamic conditions than in static conditions.

QB-10. Distinguish between the ac output power used to determine conversion efficiency and the power used in calculating power gain.

QB-11. What is one advantage and one disadvantage of the cathode follower compared with the grounded-cathode amplifier?

QB-12. How is the self-bias circuit of a cathode follower modified to provide more symmetrical operation?

QB-13. Why can a cathode follower accommodate a much larger input voltage than can a grounded-cathode amplifier?

QB-14. What main application does a grounded-grid amplifier have?

QB-15. Summarize the characteristics of the three amplifiers with respect to (i) voltage gain, (ii) input resistance, (iii) output resistance, (iv) phase relationship between input and output.

PROBLEMS

PB-1. If the cathode resistor in a triode amplifier is 4.7 K ohms, calculate the necessary value of bypass capacitor if the lowest frequency to be handled is:

(a) 1 kHz. (b) 100 Hz. (c) 10 Hz.

PB-2. Repeat Example B-2, but use in part (a) a bias voltage of -6 V instead of -4 V in determining the value of R_k, and use this new operating point for the rest of the problem.

PB-3. The circuit of Fig. B-5(a) is to be analyzed with the following values for the components:

R_g = 1 M ohm, R_k = 800 ohms, R_L = 15 K ohms, V_{PP} = 300 V, C_c = 0.1 μF.

The tube is a 12AU7 whose characteristics are given below.

(a) Draw the dc load line on the characteristic curves.

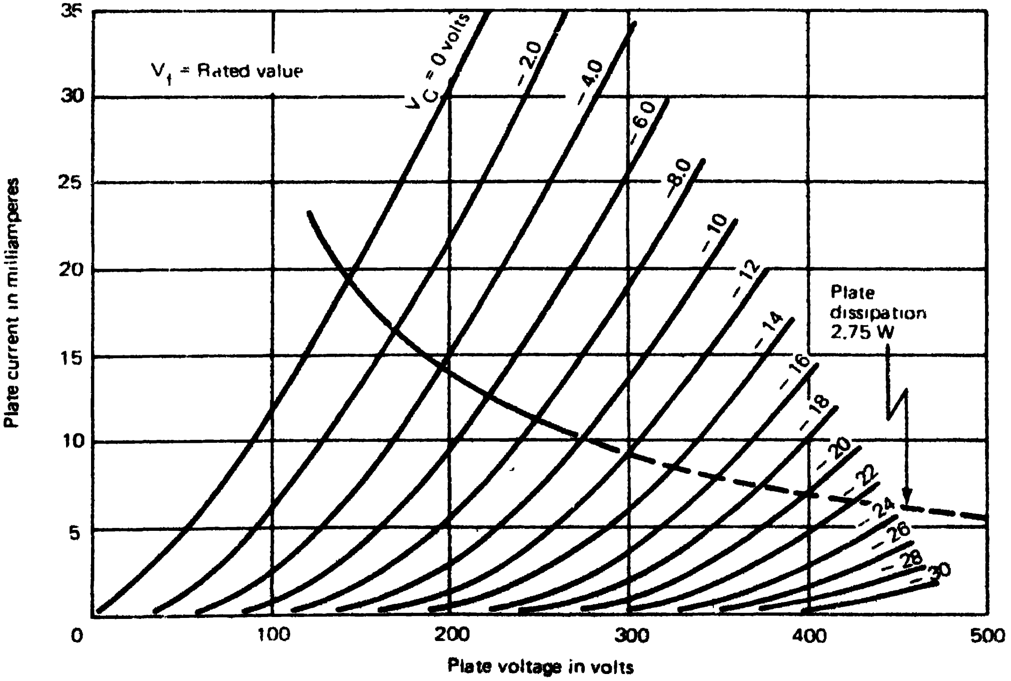
(b) Draw the bias line and determine the operating point. Specify V_P, I_P.

(c) Draw the ac load line assuming a load R_o = 20 K ohms is connected.

(d) Assuming an input signal given by V = 4 sin 2000 πt is applied, determine:

(i) The open-circuit voltage gain.
(ii) The loaded voltage gain.
(iii) The minimum value for C_k.
(iv) The percent second harmonic distortion.

6K11 (Section 1), 7AU7, 9AU7, 12AU7, 12AU7-A, 7247 (Section 2)
Average plate characteristics
each section



(Courtesy of General Electric Company)

(v) The conversion efficiency.
(vi) The plate dissipation at quiescent conditions, and when the signal is applied.
(vii) The power gain in decibels. Assume the reactance of C_c is negligible.

PB-4. Repeat part (d) of PB-3 using as the input signal the following: $V = 6 \sin 100 \pi t$. In (iv) obtain the total harmonic distortion.

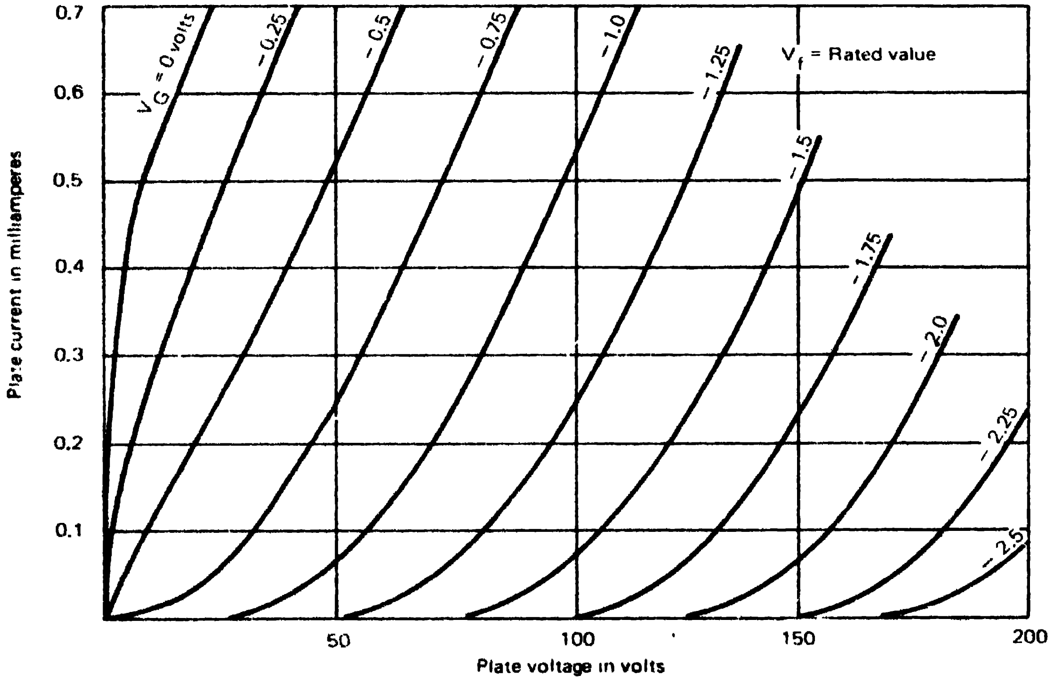
PB-5. The circuit of Fig. B-5(a) is used as a preamplifier for the magnetic phono input in an audio amplifier.

$R_g = 47$ K ohms, $R_L = 200$ K ohms and $V_{PP} = 100$ V. The tube is a 12AX7, whose characteristics are given below.

(a) Draw the dc load line on the characteristic curves.
(b) Determine the value of R_k to provide a bias of -0.75 V.
(c) Determine the voltage gain for a 1.5 V peak-to-peak input.
(d) Determine the power gain in decibels when a 500 K ohm load is connected.
(e) Determine the conversion efficiency for the conditions in (d).

PB-6. A 6AV6 has an amplification factor of 100 and a plate resistance of 80 K ohms at a plate voltage of 100 V. Calculate:

6C10, 6EU7, 6K11 (Sections 2 and 3), 12AX7, 12AX7-A, 7025, 7247 (Section 1)
Average plate characteristics
each section



(Courtesy of General Electric Company)

(a) The small-signal open circuit voltage gain of the amplifier when this tube is used in the circuit of Fig. B-8 with R_L = 100 K ohms.
(b) The voltage gain when R_o = 220 K ohms is connected, assuming C_c has negligible reactance.
(c) The voltage gain when R_o = 220 K ohms, C_c = 0.1 μF, f = 10 Hz. Also determine the phase angle between input and output.
(d) The voltage gain when R_o = 220 K ohms, C_c = 0.1 μF, f = 10 kHz, and 100 pF of capacitance exists across R_o.

PB-7. A 6BN8 with an amplification factor of 75 and a plate resistance of 21 K ohms is used in a cathode-follower circuit as in Fig. B-11(a). If R_k = 10 K ohms and R_g = 470 K ohms, calculate:

(a) The unloaded small-signal voltage gain.
(b) The output resistance.
(c) The small-signal voltage gain with a 1 K ohm load capacitively coupled to the output.
(d) The power gain in decibels for the conditions in (c).

PB-8. If the triode section of a 6BN8 were used in a grounded-grid VHF amplifier at a plate voltage of 250 V where μ = 70 and r_p = 28 K ohms, determine:

(a) The small signal voltage gain when the effective plate load resistance is 20 K ohms and the amplifier source has an internal resistance of 600 ohms.
(b) The input resistance of the amplifier.
(c) The output resistance of the amplifier.

APPENDIX C

FREQUENCY RESPONSE OF CASCADED TRIODE AMPLIFIERS

C-0. INTRODUCTION

In order to develop the gain necessary for amplifying small signals to a workable level, two or more stages may be placed in series or cascade. However, the coupling of one stage to the next has an effect on the gain at low frequencies, while the inherent capacitive effects of a tube seriously limit the higher frequency at which the triode can be operated.

The bandwidth of an amplifier will be designed to meet a certain need and determines the type of coupling used. There are four main groups of amplifiers:

(1) Direct-coupled some amplifiers must have the capability for amplifying dc and thus exclude the use of capacitors for coupling purposes.

(2) A large group of amplifiers operates in the range of frequencies from 20 Hz to 20,000 Hz and is used for audio applications. These are usually of the RC-coupled or transformer-coupled type.

(3) A third group is the video or broadband amplifier, which must handle frequencies from 20 Hz to 4 MHz and needs special compensation to achieve this capability.

(4) Radio frequency amplifiers. these normally operate with a resonant load and pass only a narrow band of frequencies.

In this appendix we will consider direct-coupled and RC-coupled amplifiers, with the emphasis on the frequency response of the latter.

C-1. DIRECT-COUPLED AMPLIFIER

In order to amplify dc voltages, as well as low frequencies, amplifier stages must be coupled directly to one another. Capacitors are obviously impractical for dc-coupling purposes.

Since the output of each stage will have some positive dc voltage level in addition to the desired signal, and since a blocking capacitor cannot be used, the next stage must have a large negative bias supply to put the operation in the middle of the characteristic curves. Such a circuit is called the Loftin-White amplifier, an example of which is shown in Fig. C-1.

By suitably arranging resistors to drop voltages, one supply voltage may be used for the V_{PP}. However, variations in the V_{PP} or bias supplies will be amplified throughout the system, causing an instability or drift problem. Consequently, other methods have been developed for dc and low frequency amplifiers; they are considered in the chapters on integrated circuits and field-effect transistors.

Assume in Fig. C-1 an input of 0.1 V peak-to-peak at a frequency of 1 Hz (as found in electrocardiography work). If the gain of each stage is 10, the output from the first stage will be 1 V peak-to-peak, the output from the second 10 V peak-to-peak, and from the third 100 V peak-to-peak. The overall gain

$$= \frac{\text{Output}}{\text{Input}} = \frac{100 \text{ V p-p}}{0.1 \text{ V p-p}} = 1000.$$

Thus total gain in 3 cascaded stages is:

$$A_T = A_1 \times A_2 \times A_3 \qquad \text{(C-1)}$$

$$= 10 \times 10 \times 10 = 1000.$$

Note that if the gain is calculated in dB the total dB gain is additive not multiplicative. That is, assume the voltages were measured across the same value of resistance, $R_i = R_o$, so that

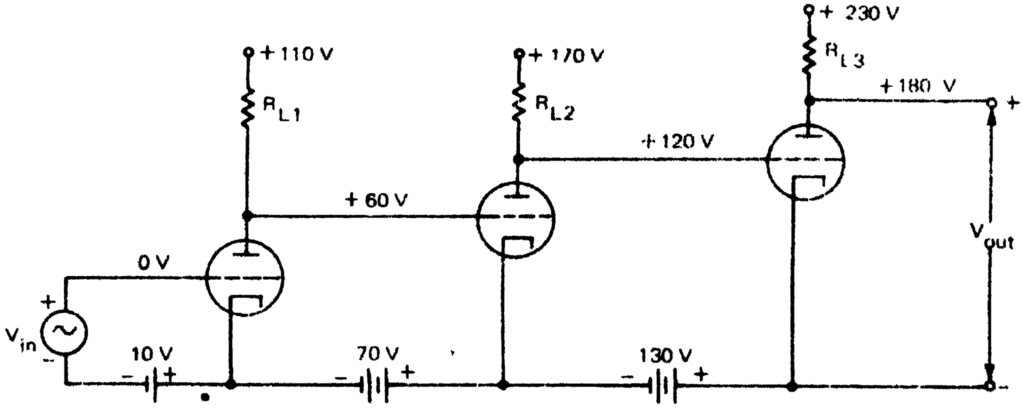


Fig. C-1. Direct-coupled Loftin-White amplifier.

$$\mathrm{dB} = 20 \log A_v. \qquad \text{(C-2)}$$

Then the dB gain per stage would be 20, but the overall gain = 20 log 1000 = 60 dB. Thus, for decibels,

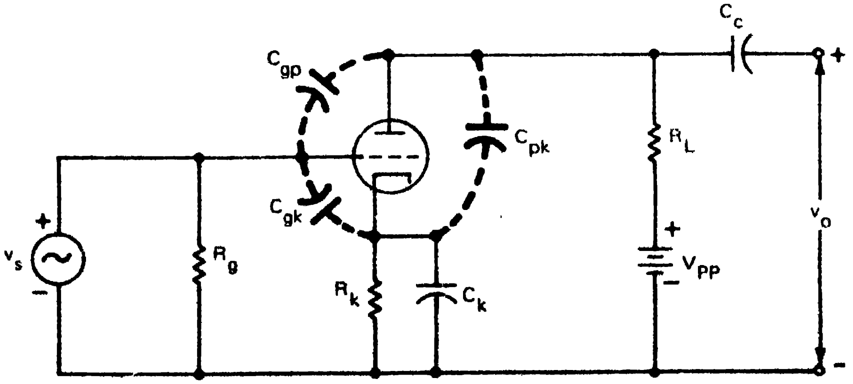
$$A_T = A_1 + A_2 + A_3. \qquad \text{(C 3)}$$

C-2. THE TRIODE AT HIGH FREQUENCIES

The limitation of the triode at high frequencies is due to the interelectrode capacitances, represented dotted in Fig. C-2.

Since C_k is much larger than the interelectrode capacitance, C_{gk} and C_{pk} are directly across the input and output respectively, as shown in Fig. C-3. Hence the effect of C_{gk} is to load the input signal, v_s, and will reduce the input signal to the grid because of an additional voltage drop across R_s, the internal resistance of the source. Therefore the output voltage will decrease.

The effect of C_{pk} is to shunt the output to ground and thus reduce v_o. Although C_{pk} is only 1.9 pF, at a frequency of 0.8 MHz this is a reactance of only 100 K ohms. Thus, if



For 1/2 12AX7:
$C_{gk} = 1.8$ pF
$C_{gp} = 1.7$ pF
$C_{pk} = 1.9$ pF

Fig. C-2. Grounded cathode amplifier showing interelectrode capacitances.

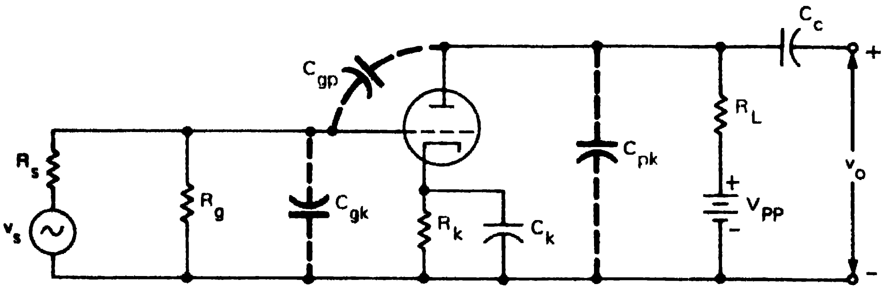


Fig. C-3. Single stage amplifier showing interelectrode capacitances.

R_L = 100 K ohms the total load impedance drops to 70 K ohms (vectorial multiplication and addition), with a corresponding decrease in voltage developed at the output. In many applications, 0.8 MHz is a relatively low frequency, so a capacitance of 1.9 pF is not negligible.

The presence of C_{gp} provides another path for input current from the source causing additional loading, but the amount of loading is surprisingly much greater than C_{gp} itself, as discussed in the next section.

C-3. TRIODE INPUT CAPACITANCE

Consider Fig. C-4, which has been drawn to represent the situation for ac signals at the input. R_g has been omitted for clarity.

Recall: V = IR or I = V/R = VG, where G = conductance. When considering reactances or, in general, impedances, the current is obtained from I = VY, where Y is the admittance.

$$\therefore I_i = V_i Y_i \text{ or } Y_i = \frac{I_i}{V_i}. \tag{C-4}$$

Let it be required to determine Y_i in terms of C_{gk} and C_{gp}.

$$I_i = I_1 + I_2. \tag{C-5}$$

$$I_1 = V_i Y_{gk}, \text{ where } Y_{gk} = \frac{1}{X_{C_{gk}}} \tag{C-6}$$

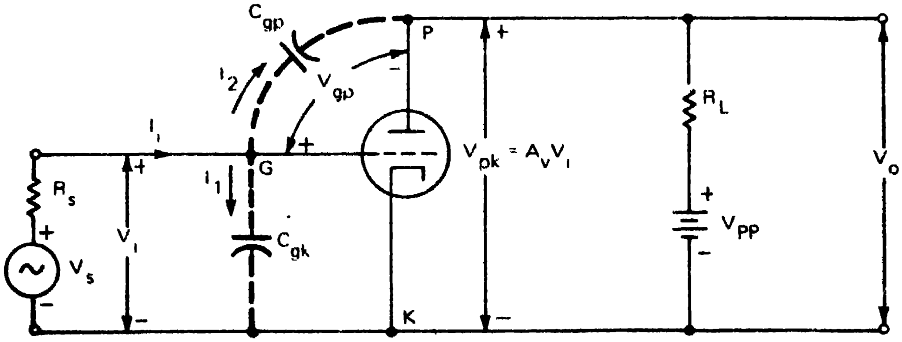


Fig. C-4. Circuit to determine triode input capacitance.

$$I_2 = V_{gp}Y_{gp}, \text{ where } Y_{gp} = \frac{1}{X_{c_{gp}}} \tag{C-7}$$

But $V_{gp} = V_i - V_{pk}$, and $V_{pk} = V_o = A_v V_i$.

$$\tag{C-8}$$

$$\therefore I_2 = (V_i - A_v V_i)Y_{gp} = V_i(1 - A_v)Y_{gp}. \tag{C-9}$$

$$\therefore I_i = I_1 + I_2 = V_i Y_{gk} + V_i(1 - A_v)Y_{gp}. \tag{C-10}$$

$$= V_i[Y_{gk} + (1 - A_v)Y_{gp}]. \tag{C-11}$$

$$\therefore Y_i = \frac{I_i}{V_i} = \frac{V_i[Y_{gk} + (1 - A_v)Y_{gp}]}{V_i} \tag{C-12}$$

$$\therefore Y_i = Y_{gk} + (1 - A_v)Y_{gp}. \tag{C-13}$$

C-4. MILLER EFFECT

Assuming that A_v is a negative real number,

$$A_v = \frac{-\mu R_L}{r_p + R_L},$$ the input admittance

(ignoring R_g) is purely capacitive, and the total input capacitance is given by

$$C_i = C_{gk} + (1 - A_v)C_{gp}. \tag{C-14}$$

This increase in input capacitance over C_{gk} by an amount $(1 - A_v)$ times C_{gp} is termed the Miller effect. The total input and output capacitances are shown in Fig. C-5.

Note that C_{gp} is no longer shown from plate to grid. It has been replaced by $C_{gp}(1 - A_v)$ connected directly across the input, and by C_{gp} across the output.

$$C_i = C_{gk} + (1 - A_v)C_{gp}. \tag{C-15}$$

$$C_o = C_{pk} + C_{gp}. \tag{C-16}$$

The Miller effect increase in capacitance may cause considerable loading on the source V_s, depending upon the frequency. This is especially true when an amplifier utilizes a high-mu tube, as shown by the following example.

Example C-1. Consider a single-stage triode amplifier with a 12AX7 high-mu tube having $\mu = 100$ and $r_p = 70$ K ohms, and working into a plate load resistance of 100 K ohms; $C_{gk} = 1.8$ pF, $C_{gp} = 1.7$ pF, $C_{pk} = 1.9$ pF. Calculate the gain of the stage at a frequency of 20 kHz, and the total input and output capacitances taking into account the Miller effect.

Solution

Output capacitance, $C_o = C_{pk} + C_{gp} = 1.9 + 1.7 = 3.6$ pF. At a frequency of 20 kHz,

$$X_{C_o} = \frac{1}{2\pi f C_o} \quad 2\pi \times 2 \times 10^4 \times 3.6 \times 10^{-12}$$

$$= 2.2 \text{ M ohms.}$$

Thus, the plate load resistance of 100 K ohms is not affected by the output capacitance, and the equation for gain developed earlier may be used.

$$A_v = \frac{-\mu R_L}{r_p + R_L} = \frac{-100 \times 100 \times 10^3}{70 \times 10^3 + 100 \times 10^3} = -59.$$

$$C_i = C_{gk} + (1 - A_v)C_{gp} = 1.8 + (1 + 59)1.7 \text{ pF}$$

$$= 103.8 \text{ pF.}$$

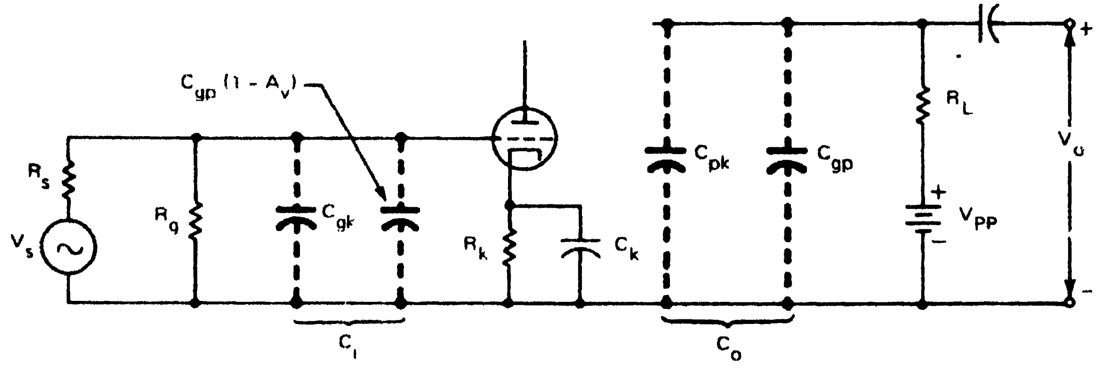


Fig. C-5. Triode amplifier showing input and output capacitances.

At a frequency of 20 kHz, X_C.

$$2\pi fC_i \quad 2\pi \times 2 \times 10^4 \times 103.8 \times 10^{-12}$$

$$= 77 \text{ K ohms}.$$

This reactance of 77 K ohms is in parallel with R_g, typically 470 K ohms, so there is a considerable increase in input current, with the total effect depending upon the relative value of R_s.

It was shown in Example C-1 that the effect of the additional input capacitance is not particularly distressing if R_s is relatively small in value. However, voltage gain is directly affected by this input capacitance when the output of one tube is coupled to the input of the next tube, as shown in the following section.

C-5. RESISTANCE-CAPACITANCE (RC) COUPLED AMPLIFIER

When it is not required to amplify dc levels, the simplest method of coupling stages is to use a capacitor, which with the grid resistor of the next stage gives RC coupling. See Fig. C-6.

It is to be noted that each stage has associated with it an input capacitance C_i and output capacitance C_o. To this should be added the stray wiring capacitance, C_s, which inevitably exists in any installation.

It is required to determine the effect of the coupling capacitor C_{c1} and the input capacitance to the next stage, C_{i2}, as well as the output capacitance C_{o1}. (Note that C_{i1} is not considered, since it would be a load on the previous stage or source and would enter into calculations for the earlier stage. Similarly C_{o2} would be part of the next stage.)

The equivalent circuit showing the resistors and capacitances involved is drawn in Fig. C-7.

Instead of obtaining a general equation of gain, (V_1/V_s), for all frequencies the amplifier stage is considered in three different regions of frequency.

Mid Frequencies

At the mid-frequency range, say 300 to 3,000 Hz (normal voice range), the shunting effect of the interelectrode and wiring capacitances is considered negligible (too low a frequency), whereas the frequency is considered high enough so that the reactance of C_{c1} is negligible compared with R_g.

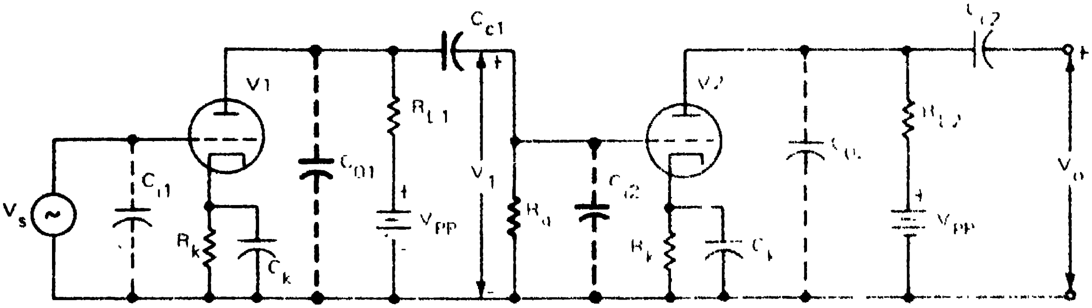


Fig. C-6. Two-stage RC-coupled amplifier showing input and output capacitances

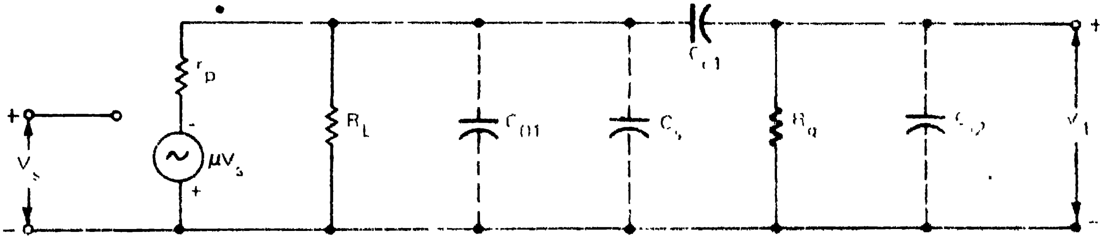


Fig. C-7 Equivalent circuit of the first stage in an RC-coupled amplifier.

Thus:

$$\frac{V_1}{V_s} = A_{vMF} = -g_m R_{eq}, \qquad \text{(C-17)}$$

where

$$\frac{1}{R_{eq}} = \frac{1}{r_p} + \frac{1}{R_L} + \frac{1}{R_g} \approx \frac{1}{r_p} + \frac{1}{R_L} \qquad \text{(C-18)}$$

Low Frequencies

At low frequencies, the shunting capacitances have negligible effect, and loss in gain is due to the coupling capacitor C_c, whose reactance becomes large compared with R_g.

It may be shown that

$$\frac{V_1}{V_s} = A_{vLF} = \frac{A_{vMF}}{1 - j\left(\frac{X_{C_c}}{R}\right)}. \qquad \text{(C-19)}$$

The lower 3-dB frequency is

$$f_1 = \frac{1}{2\pi R C_c}, \qquad \text{(C-20)}$$

where

$$R = \frac{r_p R_L}{r_p + R_L} + R_g \approx R_g. \qquad \text{(C-21)}$$

High Frequencies

The coupling capacitor C_c at high frequencies has negligible reactance compared with R_g, but the interelectrode capacitances are significant.

It may be shown that

$$\frac{V_1}{V_s} = A_{vHF} \quad \frac{A_{vMF}}{1 + j\left(\frac{R_{eq}}{X_{C_T}}\right)}. \qquad \text{(C-22)}$$

The upper 3-dB frequency is

$$f_2 = \frac{1}{2\pi R_{eq} C_T} \qquad \text{(C-23)}$$

where

$$C_T = C_{o1} + C_s + C_{i2}. \qquad \text{(C-24)}$$

C-6. FREQUENCY RESPONSE CURVE

The typical overall effect of C_c and C_T on the frequency response curve is indicated in Fig. C-8.

The preceding equations and discussion have assumed that it is economically possible to make C_k, the cathode bypass capacitor, large enough to avoid degenerative effects. This is usually the case.

C-7. GAIN-BANDWIDTH PRODUCT

Let us determine the relationship between the gain of an amplifier and its bandwidth. Does one restrict the other? That is, is it possible to increase the bandwidth of an amplifier without sacrificing gain?

The mid-frequency gain was determined as

$$A_{vMF} = -g_m R_{eq}, \qquad \text{(C-17)}$$

and the upper 3-dB frequency as

$$f_2 = \frac{1}{2\pi R_{eq} C_T}. \qquad \text{(C-23)}$$

Now the bandwidth, Δf, is $f_2 - f_1$. But f_1 is usually very small, so to a good approximation: $\Delta f \approx f_2$.

Thus

$$A_{vMF}\,\Delta f \approx A_{vMF}\,f_2 = \frac{g_m R_{eq}}{2\pi R_{eq} C_T}$$

$$= \frac{g_m}{2\pi C_T} \qquad \text{(C-25)}$$

This states that the product of an amplifier's gain and bandwidth is a constant, dependent

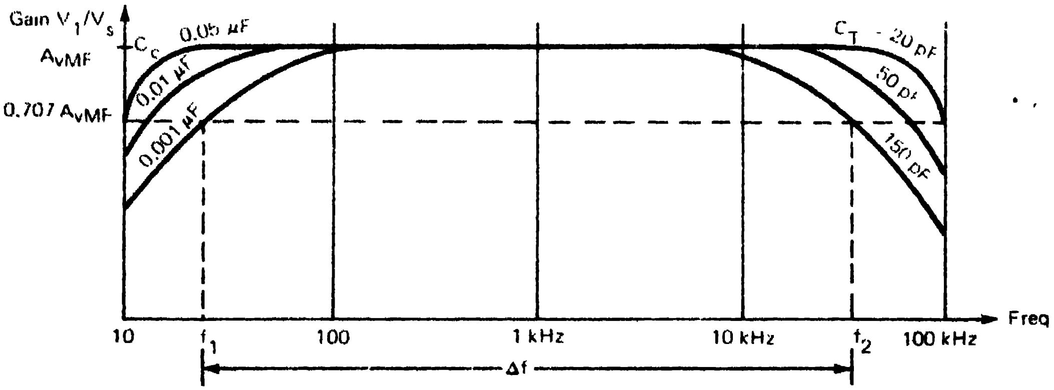


Fig. C-8. Effect of coupling and shunting capacitances on frequency response.

upon the tube's g_m, its output interelectrode capacitance, stray wiring capacitance, and the input capacitance to the next stage or load that it is driving.

For a given C_T we can see that an increase in A_{vMF} must be done at the expense of Δf and vice-versa. Put another way, if A_{vMF} is made equal to 1 (by reducing R_{eq}) then $g_m/2\pi C_T$ represents the greatest bandwidth achievable with a given tube. As gain is made to increase, the bandwidth will decrease from this value.

Let us illustrate some of the concepts of the previous three sections with a numerical example.

Example C-2. Consider the single stage amplifier in Example C-1 to be coupled to an identical amplifier, as shown in Fig. C-9.

Calculate:

(a) The mid-band gain of the first stage, V_1/V_s.

(b) The gain of the first stage, V_1/V_s at 20 kHz.

(c) The overall voltage gain, V_2/V_s at 20 kHz.

(d) The lower and upper 3-dB points for the first stage, f_1 and f_2, and hence the bandwidth.

(e) The gain-bandwidth product for the first stage using the tube's g_m and compare with the actual product using the results of (a) and (d).

$\mu = 100$, $r_p = 70$ K ohms, $R_{L1} = R_{L2} = 100$ K ohms, $C_{c1} = 0.05\ \mu$F.

Solution

(a) First Stage Mid Band Gain. At the mid frequencies C_{c1} is a short, the interelectrode capacitances are insignificant, and $A_{vMF} = -g_m R_{eq}$ (from Section C-5). Since $R_g = 470$ K ohms is much larger than $r_p = 70$ K ohms or $R_L = 100$ K ohms, this equation reverts to that used in Example C-1,

$$A_{vMF} = \frac{-\mu R_L}{r_p + R_L} = -59.$$

(b) First Stage Gain at 20 kHz. It was shown in Example C-1 that at this frequency the output capacitance C_{o2} had a reactance of 2.2 M ohms and did not affect the gain of the amplifier's second stage. Thus $A_v = -59$ for the second stage even at 20 kHz. Consequently, the input capacitance to the second stage (which is across the load of the first stage) is the value calculated in Example C-1, $C_{i2} = 103.8$ pF.

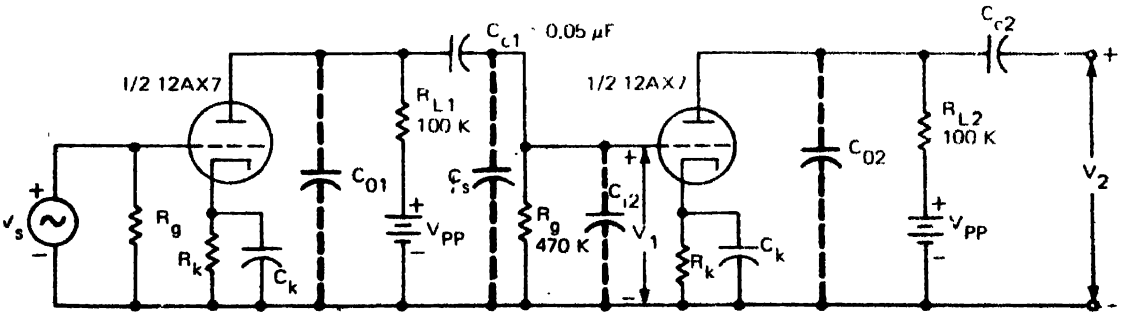


Fig. C-9. Cascaded amplifier for Example C-2.

The total shunting capacitance is

$$C_T \quad C_{o1} + C_s + C_{i2}.$$

C_{o1} was shown to be 3.6 pF, but what is C_s? If in wiring the amplifier there is a stray capacitance of 1 pF between the wires leading to the grid and plate — external to the tube — this acts as if it were part of the grid-plate interelectrode capacitance and is multiplied by the factor $(1 - A_v)$; the product in this case amounts to 60 pF Let us use this as a conservative figure representing the stray capacitance.

$$\therefore C_T = 3.6 + 60 + 103.8 = 167.4 \text{ pF}.$$

We now require the reactance of this at 20 kHz.

$$X_C = 2\pi f C_T \quad 2\pi \times 2 \times 10^4 \times 167 \times 10^{-12}$$

$$= 47.6 \text{ K ohms}.$$

Since $R_g >> r_p, R_L$ then $R_{eq} = r_p \| R_L$

$$= \frac{70 \text{ K} \times 100 \text{ K}}{170 \text{ K}} = 41.2 \text{ K ohms}.$$

From Eq. C-22

$$A_{vHF} = \frac{A_{vMF}}{1 + j\left(\frac{R_{eq}}{X_{C_T}}\right)} = \frac{V_1}{V_s}.$$

$$\therefore A_{vHF} \quad \frac{-59}{1 + j\,\frac{41.2}{47.6}} \quad \frac{59\angle 180^\circ}{1 + j\,0.87} = \frac{59\angle 180^\circ}{1.32\angle 41^\circ}$$

$$= 45\angle 139^\circ.$$

Thus, the effect of the interelectrode capacitances at high frequencies was to reduce the gain from 59 to 45 and change the phase relationship from 180° to 139°.

(c) Overall Voltage Gain V_2/V_s at 20 kHz

$$A_T = \frac{V_2}{V_1} \times \frac{V_1}{V_s} = -59 \times 45\angle 139^\circ = 2650\angle 319^\circ,$$

or an absolute value of gain = 2650, with the output 41° out of phase with the input. This compares with an overall voltage gain at the mid frequencies of $-59 \times -59 = 3480$ with no phase difference.

(d) Bandwidth of First Stage. From Section C-6,

$$f_2 = \frac{1}{2\pi R_{eq} C_T}. \qquad \text{(C-23)}$$

But $R_{eq} = 41.2$ K ohms, $C_T = 167$ pF.

$$\therefore f_2 = \frac{1}{2\pi \times 41 \times 10^3 \times 167 \times 10^{-12}}$$

$$= 23.2 \text{ kHz}.$$

This value seems reasonable, since at 20 kHz the gain had dropped to 45, while at the upper 3-dB frequency the gain should drop to $59 \times 0.707 = 41.6$.

$$f_1 = \frac{1}{2\pi R C_c} \quad 2\pi R_g C_c$$

$$2\pi \times 470 \times 10^3 \times 0.05 \times 10^{-6}$$

$$\therefore f_1 \approx 7 \text{ Hz}.$$

$$\text{Bandwidth} = \Delta f = f_2 - f_1 \approx 28 \text{ kHz}.$$

(e) Gain-Bandwidth Product of First Stage

$$A_{vMF}\,\Delta f = \frac{g_m}{2\pi C_T} = \frac{\mu}{2\pi C_T r_p}$$

$$\frac{100}{2\pi \times 167 \times 10^{-12} \times 70 \times 10^3}$$

$$= 1.36 \times 10^6 \text{Hz}.$$

From (a), $|A_{vMF}| = 59$; from (d), $\Delta f = 23$ kHz.

$\therefore |A_{vMF}|\,\Delta f = 59 \times 23 \times 10^3 = 1.36 \times 10^6$ Hz.

Note that for the second stage C_T consists of only the output capacitance C_{o2} and any stray wiring capacitance, C_{s2}.

$$C_T = C_{o2} + C_{s2} = C_{pk} + C_{gp} + C_{s2} \approx 5 \text{ pF}.$$

Thus $|A_{vMF}|\Delta f = \dfrac{g_m}{2\pi C_T}$

$$= \frac{100}{2\pi \times 5 \times 10^{-12} \times 70 \times 10^3}$$

$$= 45.5 \times 10^6 \text{Hz}$$

But we know that the mid-frequency gain of the second stage is 59.

$$\therefore \Delta f = \frac{45.5 \times 10^6}{|A_{vMF}|} = \frac{45.5 \times 10^6}{59} \text{ Hz}$$
$$= 0.77 \text{ MHz}.$$

Thus, the bandwidth of the second stage = 770 kHz, and the overall bandwidth of the amplifier is evidently determined by the first stage, = 23 kHz.

Note that if the mid-frequency gain of the first stage were reduced from 59 to 50 by reducing R_L from 100 K ohms to 70 K ohms (causing a reduction in R_{eq} from 41.2 K ohms to 35 K ohms), the new bandwidth would be

$$\frac{|A_{vMF}|\Delta f}{|A_{vMF}|'} = \frac{1.36 \times 10^6}{50} = 27 \text{ kHz},$$

compared with the original value of 23 kHz. Conversely, an increase in gain from 59 to 68, by increasing R_L from 100 K ohms to 150 K ohms, would reduce the bandwidth to $1.36 \times 10^6/68 = 20$ kHz.

This clearly illustrates the meaning of the gain-bandwidth product of the amplifier and indicates that, if a high-gain, high-bandwidth amplifier is required, a tube with a high value of g_m but a low value of interelectrode capacitance must be used. This is one of the main reasons for the development of the tetrode and pentode, which are covered in the next appendix.

C-8. NEGATIVE INPUT RESISTANCE

It was seen in Example C-2 that the gain of the first stage at high frequencies is not a real number; it is a complex number which may be represented by the following:

$$A_v = A_1 + jA_2. \tag{C-26}$$

In Example C-2, $A_v = 45\underline{/139^\circ} = -34 + j\,30$.

Consider the input admittance to the first stage, ignoring R_g for the present:

$$Y_i = Y_{gk} + (1 - A_v)Y_{gp}$$
$$= j\omega C_{gk} + (1 - A_v)j\omega C_{gp}. \tag{C-27}$$

Substitute (C-26) in (C-27):

$$Y_i = j\omega C_{gk} + (1 - A_1 - jA_2)\,j\omega C_{gp} \tag{C-28}$$

$$= j\omega C_{gk} + j\omega C_{gp} - j\omega A_1 C_{gp} + \omega A_2 C_{gp} \tag{C-29}$$

$$= \omega A_2 C_{gp} + j\omega[C_{gk} + C_{gp}(1 - A_1)] \tag{C-30}$$

$$= \frac{1}{R_i} + j\omega C_i, \tag{C-31}$$

where

$$R_i = \frac{1}{\omega A_2 C_{gp}}, \tag{C-32}$$

$$C_i = C_{gk} + C_{gp}(1 - A_1). \tag{C-33}$$

Thus the grid-to-plate capacitance not only provides an additional input capacitance of $C_{gp}(1 - A_1)$ but also an additional input resistance of $1/\omega A_2 C_{gp}$, in parallel with R_g.

For the numerical values in Example C-2, this amounts to a total input capacitance $C_i = 1.8 + 1.7(1 + 34) = 61.3$ pF and an input resistance

$$R_i = \frac{1}{2\pi \times 2 \times 10^4 \times 30 \times 1.7 \times 10^{-12}}$$

$= 156$ K ohms in parallel with R_g.

Although this may cause additional loading, the important point is this. In the above example, A_2 is a positive number, but if the load on the first stage is of the correct inductive value it is possible for A_2 to be negative. This means there is a negative resistance at the input to the amplifier! Physically this states that, instead of accepting a signal and amplifying it, the "amplifier" supplies a signal, from the output through the grid-plate capacitance to the input. That is, the circuit behaves like a self-excited oscillator if the regenerative capacitive feedback is sufficiently large, and the amplifying capability is lost. This is a very real problem in high frequency communications circuits and requires that triode amplifiers be neutralized to prevent oscillations. This consists of providing an out-of-phase component to "buck out" the plate-to-grid feedback. Details may be found in most texts covering high frequency communications circuits.

C-9. FREQUENCY RESPONSE OF A CATHODE FOLLOWER

Let us briefly consider how a cathode follower behaves at high frequencies. The interelectrode capacitances are shown in Fig. C-10.

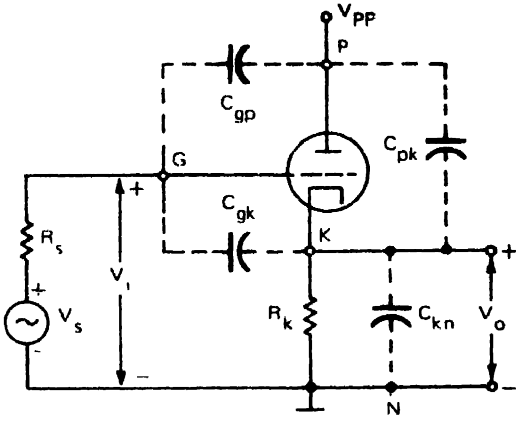


Fig. C-10. Cathode follower showing the interelectrode capacitances.

The expression for the input capacitance to the cathode follower may be obtained from calculations similar to those for the grounded cathode amplifier or by analogy with the equation for C_i of the amplifier:

For the grounded cathode amplifier, $C_i = C_{gk} + (1 - A_v)C_{gp}$. In the cathode follower, the coupling between input and output is through C_{gk} instead of C_{gp}. Also, since V_{PP} is effectively at ground level for ac, C_{gp} is directly across the input instead of C_{gk}. Hence, for the cathode follower:

$$C_i = C_{gp} + (1 - A_v)C_{gk}, \qquad \text{(C-34)}$$

where of course, A_v is positive and a little less than one.

It is clear that the less than unity gain of a cathode follower provides an extremely low value of input capacitance, which will cause minimal loading if a cathode follower is used as the second stage in a cascaded amplifier system. For a 12AX7 operating at a nominal gain of 0.9, the input capacitance $C_i = 1.7 + (1 - 0.9)1.8 \approx 2$ pF.

From a consideration of the equivalent circuit of the cathode follower and with the assumptions that $\mu + 1 \approx \mu$ and $g_m R_k >> 1$, it may be shown that the gain of the cathode follower is given by

$$A_v \quad \frac{g_m + j\omega C_{gk}}{g_m + j\omega C_T}, \tag{C-35}$$

where C_T is the total capacitance shunting R_k.

$$C_T = C_{gk} + C_{pk} + C_{kn}. \tag{C-36}$$

C_{kn} is the capacitance from cathode to ground and includes the capacitance from cathode to heater (since one side of the heater is normally grounded) as well as the output capacitance associated with the load.

Since C_T is normally much larger than C_{gk}, then approximately

$$A_v \approx \frac{}{g_m + j\omega C_T}, \tag{C-37}$$

and the decrease in gain at high frequencies is due to C_T, the upper 3-dB point occurring where $\omega C_T \quad g_m$ or,

$$f_2 \approx \frac{g_m}{2\pi C_T} \tag{C-38}$$

For a 12AX7 with a $g_m = 1.7$ millimhos, and assuming $C_T = 1.8 + 1.9 + 20 \approx 24$ pF, $f_2 \approx 11$ MHz.

It is interesting to note that the expression for f_2 (which is also the bandwidth of the circuit since the cathode follower can handle dc, so $f_1 = 0$) is the same as that developed for the gain-bandwidth expression in Section C-7. Since the gain of the cathode follower is essentially 1, then the gain-bandwidth product of a cathode follower is also $g_m/2\pi C_T$. And, since C_T is much smaller than for a grounded-cathode amplifier, the bandwidth is correspondingly much larger.

C-10. SUMMARY

(1) A direct-coupled amplifier must provide large negative bias voltages to offset the positive level associated with each output. Such an amplifier is subject to drift problems and is not very practical.

(2) The total voltage gain of an amplifier consisting of cascaded stages is the product of the voltage gains of the stages, but if power gain is measured in decibels the total power gain is the sum of the stage gains.

(3) The triode is limited in its use at high frequencies by its interelectrode capacitances causing a shunting and loading effect.

(4) The Miller effect is the additional capacitance that appears at the input to the tube and is $(1 - A_v)$ times the grid-plate capacitance.

(5) The total input capacitance to a grounded-cathode amplifier is

$$C_i = C_{gk} + (1 - A_v)C_{gp}$$

and is significant in reducing gain at high frequencies in RC-coupled cascaded amplifiers.

(6) The output capacitance of a grounded cathode amplifier is

$$C_o = C_{pk} + C_{gp}.$$

(7) The behavior of a cascaded RC-coupled amplifier is examined in three distinct frequency regions and the gain of the first stage is given by the following:

(a) Mid frequencies: $A_{vMF} = -g_m R_{eq}$

(b) Low frequencies: $A_{vLF} \quad \dfrac{A_{vMF}}{-j\left(\dfrac{X_{C_c}}{R}\right)}$

(c) High frequencies: $A_{vHF} \quad \dfrac{A_{vMF}}{1 + j\left(\dfrac{R_{eq}}{X_{C_T}}\right)}$

where $R_{eq} = r_p \| R_L \| R_g$, $R = R_g' + r_p \| R_L$,

$X_{C_c} = \frac{1}{2\pi f C_c}$, $X_{C_T} = \frac{1}{2\pi f C_T}$, and

$C_T = C_{o1} + C_s + C_{i2}$.

(8) The lower 3-dB frequency for the first stage of an RC cascaded amplifier is given by

$f_1 = \frac{1}{2\pi R C_c}$, and the upper 3-dB frequency by

$$f_2 = \frac{1}{2\pi R_{eq} C_T}.$$

(9) The gain-bandwidth product of a single stage amplifier is

$$A_{vMF}\, \Delta f = \frac{g_m}{2\pi C_T},$$

which shows that the product of gain and bandwidth is a constant for a given tube and wiring installation.

(10) The gain of an amplifier at high frequencies may be a complex number with a negative imaginary component A_2, causing the input resistance to the tube to be negative with a possibility of oscillations occurring:

$$R = \frac{1}{\omega A_2 C_{gp}}$$

(11) A cathode follower's input capacitance is much smaller than that for a grounded cathode amplifier and is given by

$$C_i = C_{gp} + (1 - A_v) C_{gk}.$$

(12) If $\mu + 1 \approx \mu$, and if $g_m R_k \gg 1$, the gain of a cathode follower at high frequencies is given by

$$A_v \approx \frac{g_m + j\omega C_{gk}}{g_m + j\omega C_T} \approx \frac{g_m}{g_m + j\omega C_T},$$

where $C_T = C_{gk} + C_{pk} + C_{kn}$.

(13) The upper 3-dB point (and bandwidth) for a cathode follower is given by

$$f_0 = \frac{g_m}{2\pi C_T}.$$

QUESTIONS

QC-1. What are the four common groups of amplifiers?

QC-2. What is a Loftin-White amplifier?

QC-3. Under what conditions is the total voltage gain of a cascaded amplifier system additive?

QC-4. What are the interelectrode capacitances of a triode due to, physically?

QC-5. What is the Miller effect? What two capacitances take the place of the grid-to-plate interelectrode capacitance?

QC-6. What are the three different regions of frequency used in analyzing the gain of an RC-coupled amplifier? In what range are the reactances of coupling and interelectrode capacitances negligible? What is the equation for voltage gain in this range?

QC-7. What is the loss in gain due to at low frequencies in an RC-coupled cascaded amplifier? When is the gain 70.7% of the value at mid frequencies?

QC-8. What is meant by the lower 3-dB point, f_1? What is the phase angle of the voltage gain at f_1?

QC-9. What causes loss in gain at high frequencies in an RC-coupled amplifier? What would be more detrimental, as far as loss in gain is concerned, a stray capacitance of 1 pF from grid to plate or 10 pF from grid to ground? Explain.

QC-10. What conditions exist in an RC-coupled amplifier stage at f_2, as far as gain and phase are concerned?

QC-11. Sketch a typical frequency response curve for an RC-coupled stage showing the relative bandwidths for the following two conditions:

(a) $C_T = 10$ pF, $C_c = 1\ \mu$F.
(b) $C_T = 200$ pF, $C_c = 0.001\ \mu$F.

QC-12. Explain what you understand by the term "gain-bandwidth product."

QC-13. If an amplifier has a gain of unity (one), what determines its maximum bandwidth?

QC-14. Explain how the grid-to-plate interelectrode capacitance can provide a negative input resistance to the tube and what effect this has on the amplifying capability.

QC-15. Why is the input capacitance to a cathode follower much smaller than in a grounded cathode amplifier?

QC-16. How does the bandwidth of a cathode follower compare with that for a grounded-cathode circuit? Why?

PROBLEMS

PC-1. Three amplifier stages are working in cascade with a 0.05 V peak-to-peak input providing a 150 V p-p output. If the voltage gain of the first stage is 20 and the input to the third stage is 15 V p-p, determine:

(a) The overall voltage gain.
(b) The voltage gain of the second and third stages.
(c) The input voltage to the second stage.
(d) The overall voltage gain in decibels and the gain of each stage in decibels using dB = 20 log A_v. Check that $A_T = A_1 + A_2 + A_3$.

PC-2. The triode section of a 6AV6 is used in an amplifier like that in Fig. C-2. $R_g = 470$ K ohms, $R_L = 100$ K ohms, $r_p = 80$ K ohms, and $\mu = 100$; $C_{gk} = 2.2$ pF, $C_{gp} = 2$ pF, $C_{pk} = 0.8$ pF.

(a) What is the output capacitance? What is its reactance at 10 kHz?
(b) Determine the open circuit voltage gain at 10 kHz.
(c) Calculate the total input capacitance, C_i, taking into account the Miller effect, and determine its reactance at 10 kHz.
(d) Calculate the total input impedance due to R_g and X_{Ci} in parallel.
(e) If a 10 kHz signal is applied that has an open circuit voltage of 0.1 V and an internal resistance of 100 K ohms, determine the magnitude and phase of the input signal to the amplifier and the output voltage from the amplifier, also in phase and magnitude. (Assume the phase angle of the input signal, V_s, is 0°.)

PC-3. Repeat Problem C-2 using a frequency of 1 MHz throughout instead of 10 kHz. What are the *two* causes for loss in gain at this frequency?

PC-4. The total input capacitance to an amplifier is determined to be 82 pF. If the amplifier's voltage gain is −35, determine the grid-to-plate interelectrode capacitance, assuming it is equal to the grid-to-cathode capacitance.

In Problems 5 through 12, refer to the circuit of Fig. C-9, where a 6FQ7/6CG7 medium-mu twin triode is used in place of the 12AX7. $C_{gp} = 3.6$ pF, $C_{gk} = 2.4$ pF, $C_{pk} = 0.34$ pF, $\mu = 20$, $r_p = 7.7$ K ohms, $R_{L1} = R_{L2} = 47$ K ohms, $R_g = 100$ K ohms, $R_k = 1860$ ohms, $C_k = 2.9$ pF, $C_{c1} = C_{c2} = 0.055\ \mu$F, $V_{PP} = 180$ V.

PC-5. Determine the mid-frequency voltage gains of:

(a) the first stage, V_1/V_s.
(b) the second stage, V_2/V_1.
(c) the overall amplifier with no additional load, V_2/V_s.

PC-6. (a) Calculate the lower 3-dB frequency, f_1, for the first stage, and

(b) the magnitude and phase of the gain at a frequency 10 Hz lower than f_1.

(c) What would be the theoretical value of f_1 for the second stage with no additional load? What then would the overall lower 3-dB frequency be for the amplifier?

(d) At what frequency will the output voltage from the first stage have a phase shift of 260° with respect to the input?

PC-7. (a) Calculate the upper 3-dB frequency, f_2, for the second stage, assuming a stray capacitance from plate-to-ground of 2 pF.

(b) Calculate f_2 for the first stage, assuming stray capacitance from plate-to-ground amounts to 2 pF, and an additional stray capacitance of 0.5 pF exists external to the grid-to-plate connections of the second stage.

(c) Determine the magnitude and phase of the gain, V_1/V_s, at a frequency 100 kHz higher than f_2 for the first stage.

(d) What would you expect the overall value of f_2 for the whole amplifier to be? What then is the bandwidth of the whole amplifier?

(e) At what frequency will the output voltage from the first stage have a phase shift of 100° with respect to the input?

PC-8. (a) What will be the overall voltage gain in magnitude and phase at 250 kHz?

(b) What voltage gain is this in dB? How does this compare with the dB gain at mid frequencies?

PC-9. (a) Determine the theoretical gain-bandwidth product of the first stage using Eq. C-25.

(b) Repeat (a) for the second stage, with stray capacitance as in PC-7(a).

(c) Determine the gain-bandwidth products for each stage using the mid-frequency gains from PC-5 and the values of f_2 from PC-7. Compare with (a) and (b).

PC-10. (a) Using the gain obtained for V_1/V_s in PC-7(c), determine the real and imaginary parts as in Eq. C-26.

(b) From the results of (a), determine the input capacitance to the first stage and the additional input resistance in parallel with R_g for the first stage, using Eq. C-32 and C-33.

(c) What total input impedance would this represent at the frequency used in PC-7(c)?

PC-11. If the second stage in Fig. C-9 is removed and replaced by a partially inductive load which results in a voltage gain of $14\underline{/200^\circ}$ at f = 419 kHz, calculate the input capacitance and additional input resistance. Comment on the possible effects.

PC-12. If the second stage in Fig. C-9 is reconnected as a cathode follower as in Fig. C-10 ($V_1 = V_i$), with $R_k = 10$ K ohms and $C_{kn} = 10$ pF, determine the following:

(a) the total capacitance C_T shunting R_k.
(b) the upper 3-dB frequency, f_2, for the second stage.
(c) the voltage gain of the cathode follower at 10 MHz using Eq. C-37.
(d) the input capacitance to the cathode follower, at 10 MHz.
(e) the value of f_2 for the first stage assuming a stray capacitance of 2 pF from plate to ground of the first

stage. Compare with the answer from PC-7(b).

(f) the bandwidth of the overall amplifier.

(g) the components of the output impedance of the amplifier at 10 MHz.

(h) the overall gain of the amplifier, V_2/V_s, at 10 MHz, in magnitude and phase.

APPENDIX D

PENTODE VOLTAGE AND POWER AMPLIFIERS

D-0. INTRODUCTION

The pentode was developed initially to provide a device that could offer significant gain at higher frequencies than was possible with the triode. However, since a pentode's characteristic curves allow a much larger area of operation than those of a triode, larger swings in output voltage and current are possible. Consequently, pentodes are found almost exclusively in tube power amplifiers where more than a few hundred milliwatts are required.

In this appendix we will consider pentode voltage amplifiers and an audio power amplifier that utilizes transformer coupling. Finally, the limitations of tube amplifiers will be discussed.

D-1. THE TETRODE

It was shown that the interelectrode capacitance from grid to plate, in the order of 2 pF, was chiefly responsible for the high-frequency limit of the triode as an amplifier. In order to extend the high-frequency capability, this capacitance must be reduced, which is accomplished by *screening* the control grid from the plate by a second grid called a *screen* grid.

If this grid were grounded, the shielding effect would be excellent, but there would be a great reduction in plate current. Therefore, the screen grid is maintained at a positive potential at from one-third to full plate voltage. This is responsible for reducing C_{gp} from 2 pF to approximately 0.02 pF; however, the other interelectrode capacitances may be increased slightly. For example, a 6CY5 used in VHF tuners in television receivers has $C_{gp} = 0.03$ pF, $C_{gk} = 4.5$ pF, $C_{pk} = 3$ pF. A 6FA7 tetrode used in complex-wave generator circuits of electronic musical instruments has $C_{gp} = 0.04$ pF, $C_{gk} = 5.5$ pF, $C_{pk} = 1.8$ pF.

The symbol for a tetrode and how it is used in a voltage amplifier are shown in Fig. D-1.

Although the screen grid may operate at the same voltage as the plate, the increase in the overall plate current may not be desirable in some cases. Fig. D-1 shows how Grid No. 2 is maintained at +100 V (for example) by the dropping resistor R_2 and the screen grid current I_{G2}. When an alternating voltage is applied to the control grid, the current flowing through the tube will vary. Thus, both i_P and i_{G2} will vary. However, this would cause v_{G2} to vary due to the changing voltage drop across R_2. Therefore a screen resistor bypass capacitor must be used ($X_{C2} \leq 0.1\ R_2$ over the bandwidth of the amplifier) to bypass the ac component and keep the dc voltage of $v_{G2} = V_{G2}$ constant. It should be noted that the total dc cathode current $I_k = I_P + I_{G2}$, which must be taken into account when calculating the value of R_k. Typical characteristic curves of the tetrode are shown in Fig. D-2 for a 6CY5.

Since the screen grid draws current it must be able to dissipate heat. The maximum input power to Grid No. 2 for the 6CY5 is 0.5 watt.

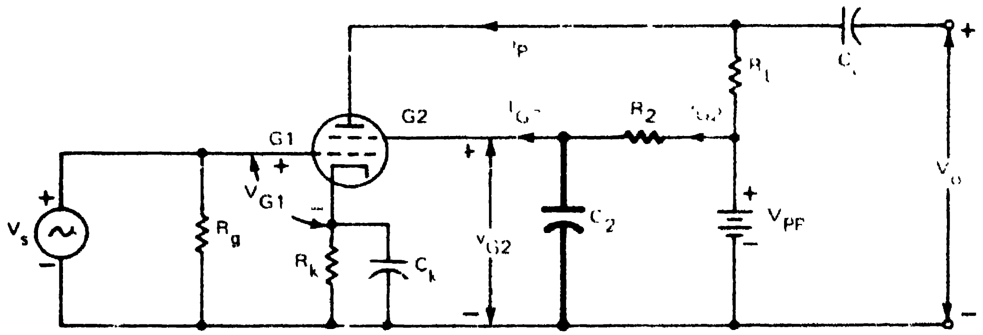


Fig. D-1. Circuit of a tetrode voltage amplifier showing screen grid G2 and dropping resistor R_2

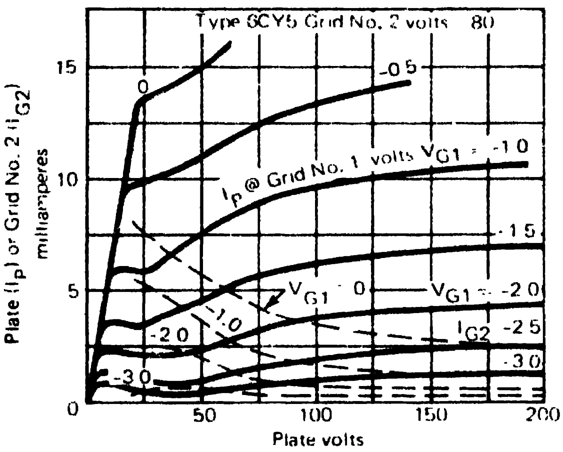


Fig. D-2. Characteristic curves of a 6CY5 tetrode. (Courtesy of RCA Corporation.)

Nonlinearities exist in the plate characteristics when the plate voltage is lower than the screen grid voltage. This is because of secondary emission of electrons from the plate which are collected by Grid No. 2, causing an increase in screen grid current at the expense of plate current. The area of operation of the tetrode is then limited to the region where $v_P > V_{G2}$. The flatter plate curves mean that the plate resistance is higher than for a triode, $r_p = \Delta v_P/\Delta i_P = 100$ K ohms for the 6CY5 and 130 K ohms for the 6FA7.

D-2. THE PENTODE

The "kink" in the plate characteristics of a tetrode can be removed by adding a third grid, Grid No. 3, called the suppressor grid. This is a very coarse grid between the screen and plate. See the symbol for the pentode in Fig. D-3 and the cutaway picture in Fig. D-4.

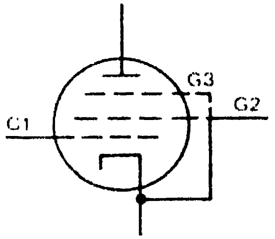
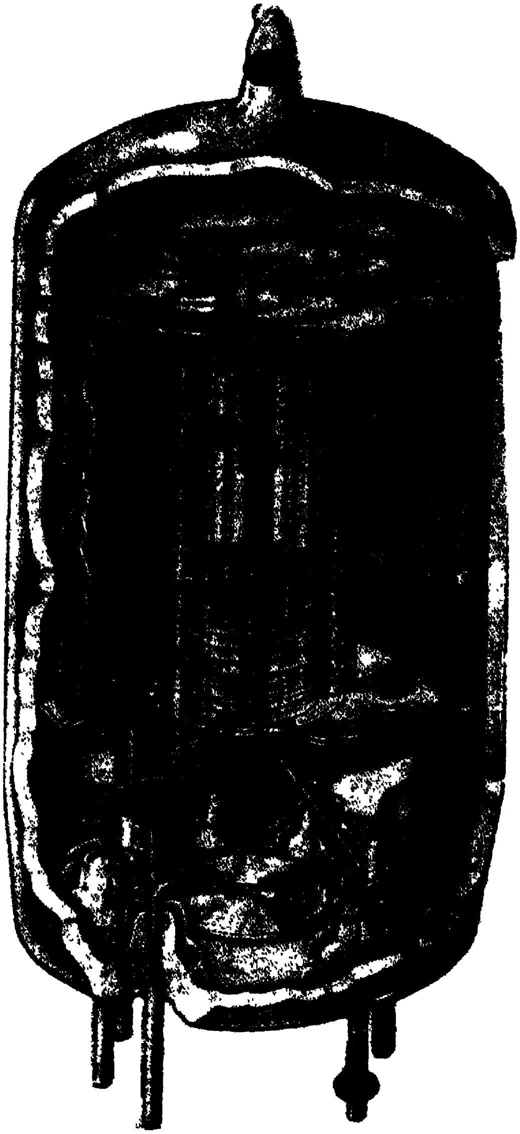


Fig. D-3. Symbol for the pentode, showing control grid G1, screen grid G2 and suppressor grid G3.

The suppressor grid is usually tied to the cathode which is much less positive than the screen grid. (This connection may be made internally in the tube during manufacture, or the grid may be left free for external



1—Glass Envelope
2—Internal Shield
3—Plate
4—Grid No. 3 (Suppressor)
5—Grid No. 2 (Screen)
6—Grid No. 1 (Control Grid)
7—Cathode
8—Heater
9—Exhaust Tip
10—Getter
11—Spacer Shield Header
12—Insulating Spacer
13—Spacer Shield
14—Inter-Pin Shield
15—Glass Button-Stem Seal
16—Lead Wire
17—Base Pin
18—Glass-to-Metal Seal

Fig. D-4. Structure of a miniature pentode tube. (Courtesy of RCA Corporation.)

connection.) This has a repelling effect on the secondary electrons that are emitted from the plate, and suppresses these electrons from going to Grid No. 2, the screen grid. Hence the curves are even flatter as shown in Fig. D-5, with r_p approximately 1 M ohm for a 6AU6 sharp cutoff pentode.

As a comparison of some of the properties of triodes, tetrodes and pentodes consider Table D-1.

Triodes are referred to as being low-, medium-, or high-mu types, depending upon whether the mu is less than 10, between 10 and 50, or more than 50, respectively.

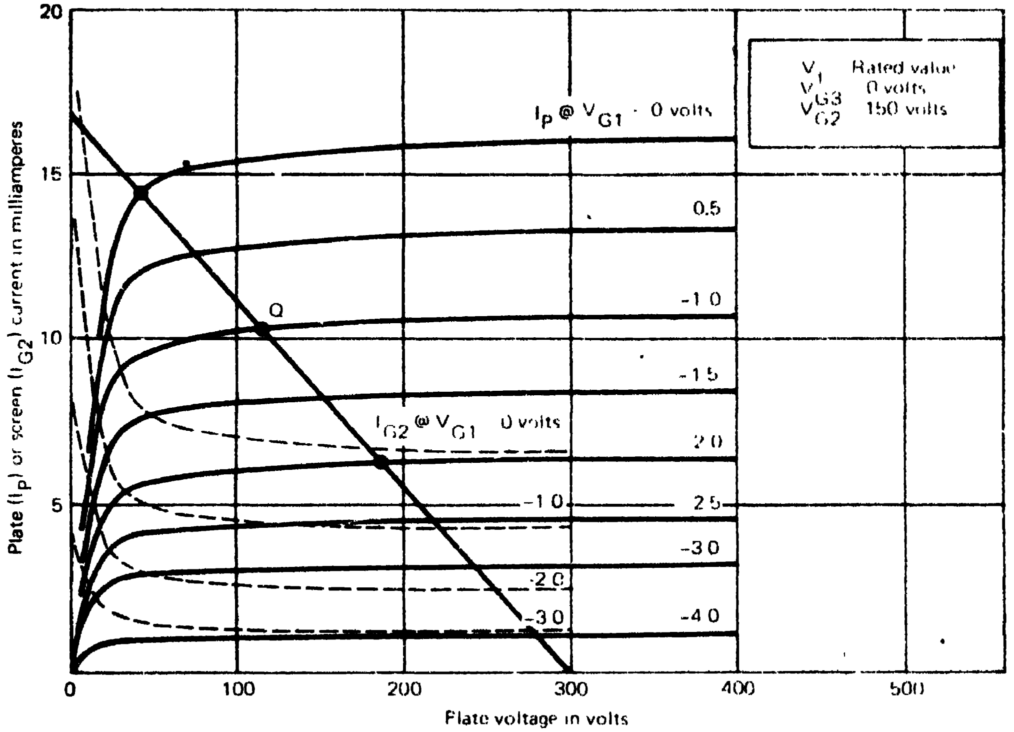


Fig. D-5. Average plate characteristics for a 6AU6, and load line for Example D-1. (Courtesy of General Electric Company.)

For a 6AU6 the C_{gp} has been reduced still further to 0.0035 pF, although this is unusually low; C_{gp} is typically nearer 0.05 pF except for power pentodes which may be as much as 0.5 pF. The grid-to-cathode and plate-to-cathode capacitances are appreciably larger in all pentodes, being as high as 10 pF.

However, tube manuals do not give the mu for tetrodes or pentodes, just the r_p and g_m. But since $\mu = r_p g_m$ as for triodes, mu can be calculated and is shown in parentheses in Table D-1.

Tetrodes and pentodes may be designated sharp-, semiremote-, or remote-cutoff types and are based on the ratio, in percent, of the

TABLE D-1 COMPARISON OF TRIODE, TETRODE, AND PENTODE PROPERTIES

	r_p, K ohms	g_m, µmhos	µ	C_{gp}, pF	C_{gk}, pF	C_{pk},
6CG7 (triode)	7.7	2600	20	4.0	2.3	2.2
6CY5 (tetrode)	100	8000	(800)	0.03	4.5	3.0
6AU6 (sharp-cutoff pentode)	1000	4500	(4500)	0.0035	5.5	5.0
6K6 (power pentode)	100	2100	(210)	0.5	5.5	6.0
6V6 (beam power tube)	50	4000	(200)	0.3	10.0	11.0

negative control-grid voltage to the screen-grid voltage, for cutoff. For sharp cutoff, the ratio is less than 10 percent, for semiremote 10 to 20 percent, and for remote more than 20 percent. A 6AU6 with a screen grid voltage of 100 V has a plate current of only 10 µA when the control grid is -4.2 V.

D-3. VOLTAGE GAIN OF A PENTODE AMPLIFIER

The same equivalent circuit may be used for the pentode amplifier as was developed for the triode. Hence, the equation for voltage gain is the same at mid frequencies:

$$\frac{-\mu R_L}{r_p + R_L} \qquad \text{(D-1)}$$

Recalling that $\mu = r_p g_m$, this becomes

$$A_v = \frac{-g_m r_p R_L}{r_p + R_L} \qquad \text{(D-2)}$$

and dividing numerator and denominator by r_p, this becomes

$$A_v = \frac{-g_m R_L}{1 + \frac{R_L}{r_p}} \qquad \text{(D-3)}$$

But in the pentode, r_p is very large and usually exceeds R_L by at least a factor of 10, so R_L/r_p becomes quite small, with the result that the voltage gain in a pentode is given approximately by

$$-g_m R_L \text{ wherever } r_p \gg R_L. \qquad \text{(D-4)}$$

For example, if a 6AU6 pentode with $r_p = 1$ M ohm and $g_m = 3900$ µmhos is used with an effective load R_L of 18 K ohms, the voltage gain will be

$$A_v = -3900 \times 10^{-6} \times 18 \times 10^3 = -70.$$

The pentode may also be used in common-plate and common-grid configurations with the same equations applying as were developed for the triode. The screen and suppressor grids must be bypassed to the cathode, not to ground.

D-4. GRAPHICAL ANALYSIS OF A PENTODE VOLTAGE AMPLIFIER

The analysis of a pentode amplifier is similar to that for a triode and is best shown by an example.

Example D-1. Given the pentode voltage amplifier in Fig. D-6, use the characteristic curves in Fig. D-5 to determine:

(a) The value of R_k to provide a bias of −1 volt.

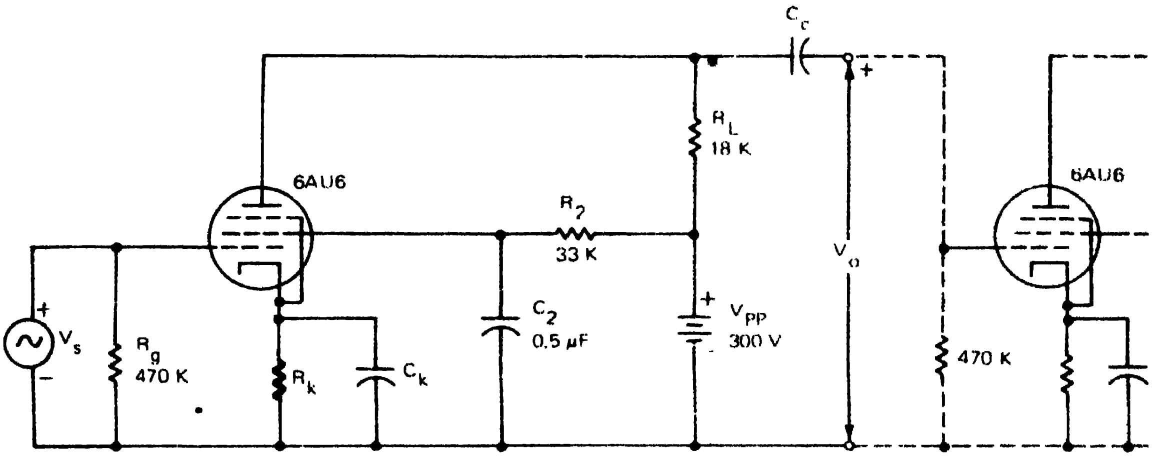


Fig. D-6. Pentode voltage amplifier for Example D-1.

(b) The peak-to-peak output voltage V_o and voltage gain for a 2 V p-p input.

Solution

(a) **Value of R_k.** Assuming R_k will be small compared with R_L, draw the dc load line and determine the Q point at the intersection with $V_{G1} = -1$ volt:

$$I_P = 10.3 \text{ mA}, \ V_P = 115 \text{ V}, \ I_{G2} = 4.5 \text{ mA}.$$

Note that I_{G2} was determined from the dotted line corresponding to $V_{G1} = -1$ V at a plate voltage of 115 V.

Total cathode current $= I_P + I_{G2} = 10.3 + 4.5 = 14.8$ mA.

$$\frac{v_k}{I_k} \qquad \overline{14.8 \times 10^{-3}} \approx 68 \text{ ohms.}$$

Note: The determination of a Q point for a given value of R_k (same conditions of V_{PP} and R_L) is not as easy and involves repeated trials to find the correct I_P and I_{G2} so that $V_{G1} = -(I_P + I_{G2})R_k$ falls on the load line.

You may wish to try finding Q for $R_k = 100$ ohms.

(b) **Voltage gain.** For a 2 V p-p input signal the grid will be driven up to 0 V and down to −2 V along the dc (and ac) load line. Corresponding plate voltage is as follows:

$$v_{G1} = 0 \text{ V}, v_p = 40 \text{ V} \qquad = 14.4 \text{ mA}$$

$$V_{G1} = -1 \text{ V}, V_P = 115 \text{ V} \ I_P = 10.3 \text{ mA}$$

$$v_{G1} = -2 \text{ V}, v_P = 190 \text{ V} \ i_P = 6.2 \text{ mA}$$

$$V_{o,p\text{-}p} = 190 - 40 = 150 \text{ V and}$$

$$|A_v| = \frac{V_o}{V_-} = \frac{150}{2} = 75,$$

which compares quite well with the calculated value in Section D-3.

D-5. HIGH-FREQUENCY RESPONSE OF A CASCADED PENTODE VOLTAGE AMPLIFIER

Assume that the voltage amplifier in Fig. D-6 has an identical pentode amplifier coupled to

its output. What would be the upper 3-dB frequency for the first stage?

To calculate f_2 we would use the same equation developed for the cascaded triode:

$$2\pi R_{eq} C_T \qquad \text{(D-5)}$$

where $R_{eq} = r_p \| R_L \| R_g = 1\ \text{M} \| 18\ \text{K} \| 470\ \text{K} \approx 18$ K ohms.

However, what is C_T? This is the parallel combination of the first pentode's output capacitance, the stray wiring capacitance, and the input capacitance to the second pentode:

$$C_T \quad C_{o1} + C_s + C_{i2}. \qquad \text{(D-6)}$$

For a pentode:

$$C_o = C_{pk} + C_{ps} + C_{p3}, \qquad \text{(D-7)}$$

where C_{ps} and C_{p3} are the plate-to-screen and suppressor grid capacitances, and are included in the output capacitance given in a tube manual. This was abbreviated to C_{pk} in Table D-1, so C_o for a 6AU6 is 5 pF.

$$C_{i2} \quad C_{gk} + C_{gp}(1 - A_v) + C_{gs}, \qquad \text{(D-8)}$$

where C_{gs} is the grid-to-screen capacitance. But since C_{gp} is only 0.0035 pF, the Miller effect is now negligible and $C_{i2} = C_{gk} + C_{gs}$. This is referred to by the tube manual as the input capacitance, and is 5.5 pF for the 6AU6. This was abbreviated to C_{gk} in Table D-1.

$\therefore C_{i2} = 5.5$ pF and $C_T = 5 + C_s + 5.5$ pF.

It is clear that the high-frequency response now depends more on the stray wiring capacitance C_s than on the tube itself. For example, if the stray wiring capacitance from grid to plate exterior to the tube is only 1 pF we have effectively made $C_{gp} = 1$ pF, and with a tube gain of 70 we have, due to the Miller effect, a C_s of at least 70 pF. Thus, the substitution of a pentode for a triode will not automatically guarantee an increase in f_2 if wiring has not conformed to good high frequency practice, namely, short leads properly dressed.

Assuming $C_s = 0$,

$$f \quad \frac{1}{2\pi \times 18 \times 10^3 \times 10.5 \times 10^{-12}}$$

$$0.84\ \text{MHz}.$$

But if $C_s = 70$ pF, $f_2 = 110$ kHz.

Note that the higher g_m and the lower (minimum) value of C_T increases the gain-bandwidth product, $g_m/2\pi C_T$ for a pentode compared with a triode (2×10^6 for a 12AX7 but 62×10^6 for a 6AU6). Thus, for the same gain, a pentode has a much higher frequency or bandwidth capability than the triode.

D-6. LOW-FREQUENCY RESPONSE OF A CASCADED PENTODE VOLTAGE AMPLIFIER

The lower 3-dB frequency, f_1, for the cascaded amplifier in Fig. D-6, is determined by both C_c and C_k. Typical values for C_c are 0.01 or 0.1 μF. C_k must be much larger, in the order of 50 μF, although if $g_m R_k << 0.414$ there may not be a lower 3-dB frequency due to C_k and R_k if C_k is properly chosen.*

D-7. RADIO-FREQUENCY (RF) VOLTAGE AMPLIFIERS

Let us briefly look at how a pentode is used as an amplifier at radio frequencies, and in particular at a frequency of 10.7 MHz, which is the intermediate frequency (IF) used in FM (frequency-modulated) radio receivers.

---

*See A. Mottershead, *Laboratory Experiments for Electronic Devices and Circuits*, (Pacific Palisades, Calif.: Goodyear Publishing Co., 1973), Experiment D.

Although a complete discussion is beyond the scope of this text, there are some important principles that should be covered.

We saw in Section D-5 that even a 6AU6 with its extremely low value of C_{gp} (0.0035 pF) had a maximum upper 3-dB point f_2 of 0.84 MHz with $R_L = 18$ K ohms. This could be increased somewhat by a reduction of R_L, but then the gain would suffer. And in an FM receiver where the typical input signal is of the order of only 20 μV, substantial amplification is necessary in the three or more RF amplifiers to provide a nominal output of 0.5 V. Since the restriction on high-frequency operation (above 1 MHz) is now caused by C_{gk} and C_{pk}, a method to nullify their effect is needed. (This amounts to an effective reduction in C_T in the expression $g_m/2\pi C_T$, which is the gain-bandwidth product.)

The answer is found in the use of tuned parallel resonant circuits to act as the load on the amplifier and also provide a means of coupling to the next stage, as shown in Fig. D-7.

It is clear from the diagram that the output capacitance of V1, $C_{o1} = C_{pk}$, is in parallel effectively with C_1, which with L_1 forms a parallel circuit. This circuit may be tuned to the resonant frequency of 10.7 MHz by the adjustment of L_1. Similarly, the input capacitance of V2, $C_{i2} = C_{gk}$, is in parallel with C_2 and L_2, which are also adjusted to resonance. T1 is the IF transformer (also affectionately referred to as the "IF can" since it is enclosed in a grounded metal shield); the voltage developed across its output, V_1, is applied to the next stage which will also have an IF transformer for load and coupling purposes.

The adjustments of L_1 and L_2 affect not only the resonant frequency but also the degree of coupling between primary and secondary. This in turn determines the frequency response or bandwidth, which in an FM receiver IF amplifier must be in the order of 200 kHz. This relatively small bandwidth allows stage gains, V_1/V_s, in the order of 100 with no difficulty at all. It should be noted in passing that, if the cascaded RF amplifiers have been aligned (tuned to resonance) with a given set

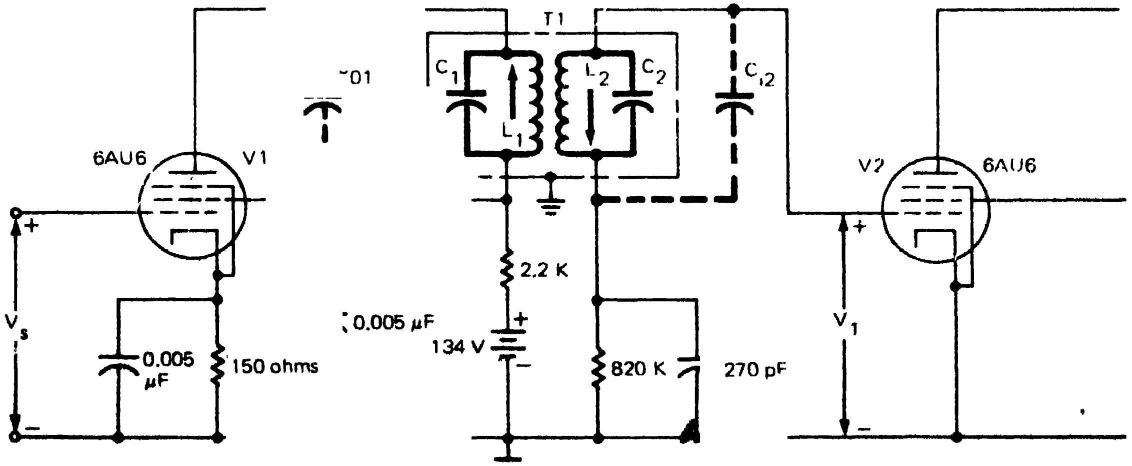


Fig. D-7. RF voltage amplifier using tuned resonant plate load.

of tubes, the replacement of a faulty tube at some later time may cause incorrect overall operation if the interelectrode capacitances are not exactly the same. Some slight realignment may be necessary.

Detailed equations relating V_1 to V_s are available in texts on communications circuits, but the gain provided by tube V1 from V_s to the primary of T1 is given by $A_v = -g_m R_e$, where R_e is the equivalent load reflected from the secondary into the primary of T1 and is resistive at resonance. More information on transformer coupling is covered in the next section.

D-8. POWER AMPLIFIERS

Power Pentode Tube

When it was realized that a pentode could provide much higher swings in plate voltage and current than a triode, emphasis was placed on enlarging the electrodes so that greater power capability could be obtained. The result was the power pentode, which is widely used in audio power amplifiers. For example, a 6K6 has a plate dissipation capability of 8.5 W, while a 7868 can dissipate 19 W. The inevitable increase in interelectrode capacitance did not cause concern since the necessary bandwidth in such amplifiers was only 20 kHz. These power pentodes showed a marked decrease in plate resistance, 100 K ohms for the 6K6 and 29 K ohms for the 7868, with corresponding increases in g_m.

Beam Power Tube

An alternative way of modifying the tetrode is to add deflection plates to concentrate the electrons into a dense narrow beam between screen grid and plate, as illustrated in Fig. D-8(a). A symbol that is sometimes used to designate such a tube is shown in Fig. D-8(b).

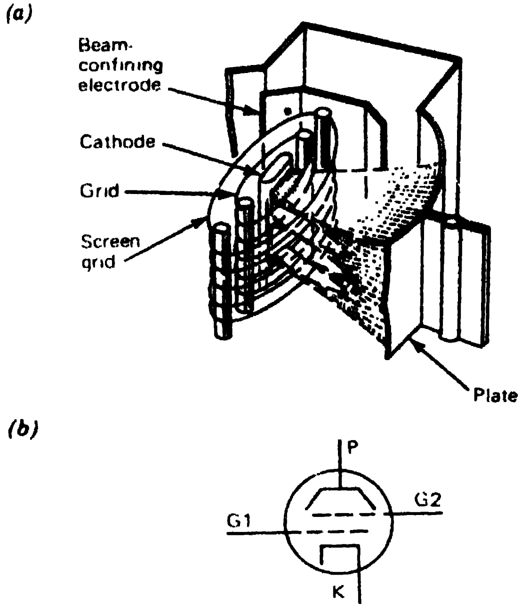
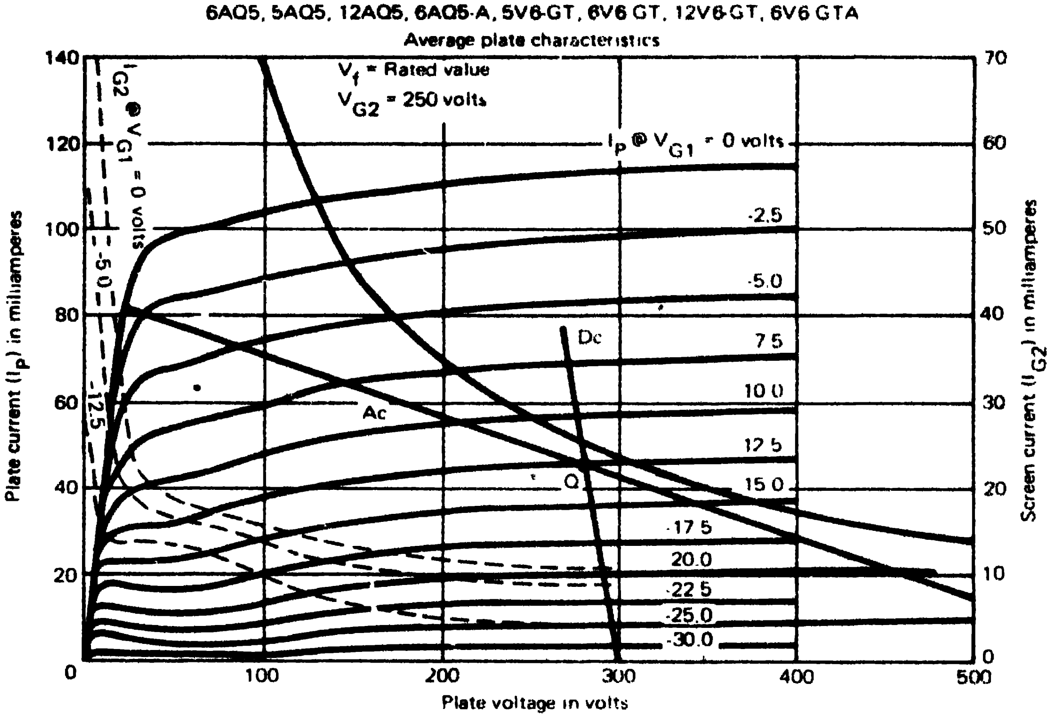


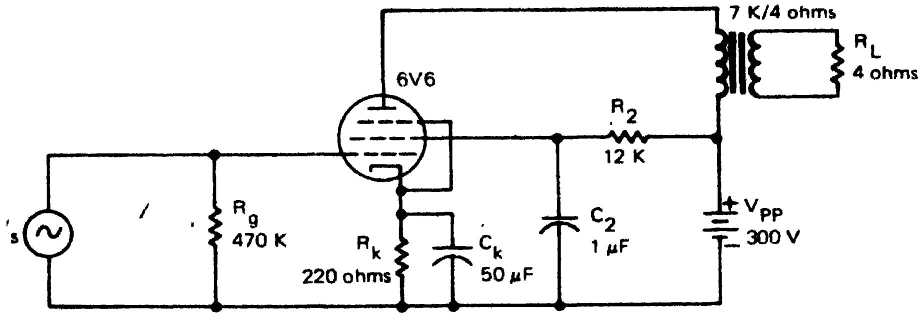
Fig. D-8. (a) Structure of beam power tube showing beam-confining action. (Courtesy of RCA Corp.) (b) Symbol sometimes used to represent a beam power tube.

However, many of the beam power tubes use a suppressor grid for beam concentration, and the symbol is the same as for the pentode.

High-power beam-power tubes such as the 8417 have a plate dissipation of 35 watts and a plate resistance as low as 16 K ohms. Typical plate characteristics for a beam power tube are shown in Fig. D-9. Note that at low plate and large negative grid voltages the "kink" associated with a tetrode tends to reappear. However, the manner in which the tube is used in transformer-coupled audio power amplifiers avoids this area of operation. This is indicated by the quiescent point located at the intersection of the dc and ac load lines in Fig. D-9. These have been drawn for the transformer-coupled power amplifier shown in Fig. D-10. The beam power tube in such an amplifier has the advantages of high power



ig. D-9. *Characteristic curves of a beam power tube showing the dc and ac load lines for the amplifier in* ig. *D-10. (Courtesy of General Electric Company.)*


7 K/4 ohms
R_L
4 ohms
6V6
R_2
12 K
R_g
470 K
R_k
220 ohms
C_k
50 μF
C_2
1 μF
V_{PP}
300 V
s

ig. *D-10. Transformer-coupled power amplifier.*

output, high power sensitivity, and high efficiency, because of low screen grid current and effective suppressor action.

Transformer-Coupling

The transformer is used to couple the load to the output of the tube as shown in Fig. D-10.

The use of a transformer provides three very important functions.

(1) The very low dc resistance of the transformer primary means that negligible dc power is lost compared with a load resistor in an RC-coupled amplifier. This means the conversion efficiency is higher, and it may be shown to have a maximum theoretical value of 50 percent.

(2) The dc load line for the amplifier is practically vertical (from V_{PP} on the plate voltage axis), since R_k and the transformer primary resistance are quite low. Thus, by proper choice of R_k, the operating point may be placed in the region of the characteristic curves near the maximum dissipation hyperbola, so as to provide a condition for maximum power output, and to avoid the "kinks" in the curves.

(3) The output transformer matches the high output impedance of the power tube to the low impedance of the load.

Although the design of such an amplifier will not be carried out here[*] the selection of the ac load line will determine both the output power and harmonic distortion. The variation of power and distortion with effective plate load resistance is indicated in Fig. D-11.

At some value of load resistance, b, the power output is maximum, but the total harmonic distortion is a minimum at some lower value of load resistance, a. The amplifier would normally be operated at "a" with a sacrifice in output power to achieve lower distortion.

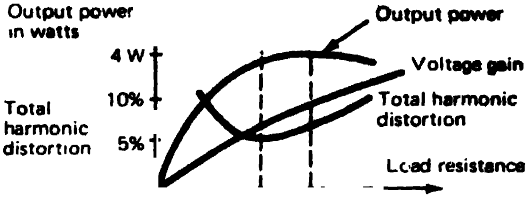


Fig. D-11. Curves showing how power output and harmonic distortion vary with load resistance for a pentode power amplifier.

Frequency Response of a Transformer-Couple Audio Power Amplifier

A typical frequency response curve for the gai of the amplifier in Fig. D-10 is shown in Fig. D-12.

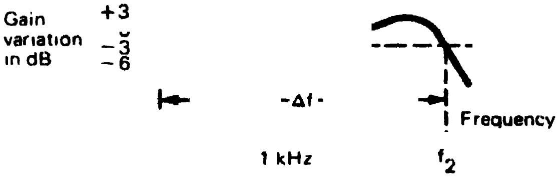


Fig. D-12. Typical frequency response curve of a transformer-coupled amplifier.

This curve shows a resonant rise that may result at higher frequencies due to the distributed capacitance and inductance of the transformer. Except for quality output transformers, the lower 3-dB frequency will be determined by the shunting effect of the transformer's primary inductance, rather than the bypass capacitors.

D-9. THE TRIODE CASCODE AMPLIFIER

As discussed in Section D-10, the pentode tends to be noisier than a triode due to the screen grid current variations. A circuit that acts effectively as a pentode but which needs no screen current is the cascode amplifier which consists of two triodes in series (same plate current). See Fig. D-13.

---

[*] See A. Mottershead, *Laboratory Experiments for Electronic Devices and Circuits*, (Pacific Palisades, Calif.: Goodyear Publishing Co., 1973), Expt. D-4.

$$A_v = \frac{V_o}{V_s} = -g_m R; I_{G2} = 0,$$

since V_{G2} is negative, $V_{P1} > V'$.

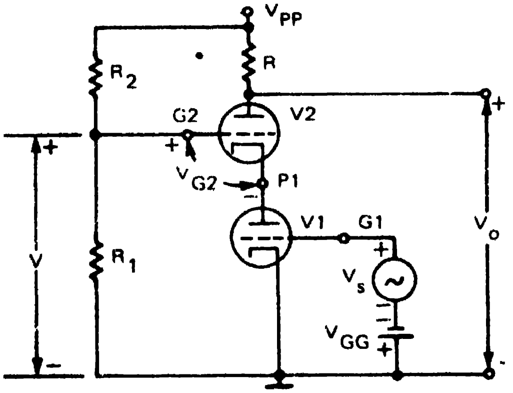


Fig. D-13. Triode cascode amplifier that acts like a pentode.

D-10. LIMITATIONS OF AMPLIFIERS USING TUBES

Voltage amplification in a cascade amplifier can be as much as a million, but the amount of gain must be considered along with the unwanted signal generated by noise currents developed in the amplifier. These "noise currents" arise from several sources:

(1) "Shot noise," generated by the random emission of electrons from the cathode.

(2) "Thermal noise," due primarily to the movement of free electrons in the grid circuit (usually in the grid-leak resistor). Noise voltage generated is given by

$$V_n = \sqrt{4KTBR}\,, \qquad \text{(D-9)}$$

where: K is the Boltzman constant $= 1.381 \times 10^{-23}$ joules/°K.
T is the temperature of the resistor in °K = °C + 273°.
B is the bandwidth over which the noise is measured in Hz.
R is the value of the resistor in ohms.

For example, if R = 1 M ohm at room temperature, and B = 20 kHz, then $V_n = 18\ \mu V$ rms.

(3) "Partition noise," present in tetrodes and pentodes due to slight changes in the number of electrons falling on the screen grid as a result of random variations in transit to the plate.

(4) Microphonics, caused by vibration of the elements within the tube or of other components and wiring in the circuit.

(5) Hum and other noises, caused by stray electrostatic or electromagnetic fields and by improperly filtered supply voltage.

The transistor overcomes some of these problems, but is not completely free of "noise." However, the superiority of transistors over the tube results from their small size, low voltage, and no need for heater supply.

D-11. SUMMARY

(1) The tetrode is a modified triode, with a screen grid added to reduce grid-to-plate interelectrode capacitance.

(2) A pentode is the result of adding a suppressor grid to a tetrode to remove the "kink" in the characteristic curves and further reduce interelectrode capacitance.

(3) Some pentodes are designed for high-frequency voltage amplification and have a

very high plate resistance, with the voltage gain given by $A_v = -g_m R_L$.

(4) Since the grid-to-plate interelectrode capacitance is practically negligible in a pentode voltage amplifier, the Miller effect no longer limits high frequency operation. However, external wiring capacitances and other interelectrode capacitances do limit f_2 in a cascaded RC amplifier.

(5) The lower 3-dB frequency f_1 may be determined by C_c or C_k or both, but if $g_m R_k \leq 0.414$ an f_1 does not exist as far as C_k is concerned.

(6) An RF voltage amplifier utilizes a tuned resonant circuit for load and coupling purposes, causing an effective reduction of C_T and allowing the amplifier to work at frequencies of 10 MHz with considerable gain.

(7) Power pentodes and beam power tubes are used as audio power amplifiers with transformer coupling, to provide a higher conversion efficiency and optimum placement of the operating point.

(8) Design of a power amplifier involves selecting an ac load line to maximize power output consistent with reasonable distortion; determination of R_k is based on the fact that total cathode current $I_k = I_P + I_{G2}$.

(9) A triode cascode amplifier consists of two triodes in series that provide a pentode's characteristics without the noise arising from screen grid current.

(10) A tube amplifier is limited by several sources of noise, a pentode being noiser than a triode due to the additional effect of "partition noise."

QUESTIONS

QD-1. Why is the screen grid of a tetrode held at a positive potential? Which capacitance is reduced by the action of the screen grid?

QD-2. What causes the nonlinearities in a tetrode's plate characteristics?

QD-3. Why is the name "suppressor grid" appropriate for the third grid in a pentode?

QD-4. What are four major effects of adding the screen and suppressor grids compared with a triode?

QD-5. What do the words "sharp, semiremote and remote" refer to when used with pentodes?

QD-6. Does the small-signal equivalent circuit for a triode apply equally well to the tetrode and pentode? What does this imply about the equations for voltage gain, etc.? What approximation can usually be made for the voltage gain of a pentode amplifier?

QD-7. The pentode was developed to reduce the input capacitance in cascaded stages, by making the Miller effect negligible. Under what conditions is the Miller effect not negligible in a pentode amplifier, and what precautions should be taken to avoid this difficulty?

QD-8. How does the gain-bandwidth product of a pentode amplifier compare with that for a triode? Explain why.

QD-9. Under what conditions does a pentode voltage amplifier with a bypassed self-bias resistor, R_k, not have a lower 3-dB frequency, f_1, due to C_k?

QD-10. Explain in detail how a radio-frequency voltage amplifier can work at frequencies as high as 10 MHz with considerable gain, if the same tube used in an RC-coupled amplifier has an upper 3-dB frequency of approximately only 1 MHz.

QD-11. What is meant by "aligning" a cascaded RF amplifier? When may realignment be necessary? What causes this?

QD-12. Why is a power amplifier often operated at a load resistance somewhat lower than that which would provide maximum power output?

QD-13. What two important functions in addition to impedance matching are provided by using transformer coupling?

QD-14. What advantage does a triode cascode amplifier have over a pentode amplifier? Why?

QD-15. Which would you expect to be noisier, a triode or a pentode? Why?

PROBLEMS

PD-1. A 6L6 beam power tube is operated at a plate voltage of 250 V in a class A amplifier, with $r_p = 22.5$ K ohms and $g_m = 6000$ μmhos. If the amplifier plate load resistance $R_L = 2.5$ K ohms, calculate the voltage gain using:

(a) the approximate equation (D-4).
(b) the more exact equation (D-3).

PD-2. Repeat PD-1 with $R_L = 5$ K ohms. What can you conclude regarding the use of the approximate equation?

For Problems 3 through 5 refer to Fig. D-6 and the characteristic curves of the 6AU6 in Fig. D-5.

PD-3. (a) Determine the value of R_k to provide a bias of −1 volt for the following conditions:

(i) $V_{PP} = 300$ V, $R_L = 20$ K ohms.
(ii) $V_{PP} = 300$ V, $R_L = 15$ K ohms.

(b) For the two load conditions above, determine the peak-to-peak output voltage V_o and the voltage gain for a 2 V p-p input. Determine if the output voltage and current swing equally from the quiescent values; compare with the result in Example D-1. What effect would the above two loads have on the second harmonic distortion compared with $R_L = 18$ K ohms?

PD-4. Given $R_L = 18$ K ohms and $V_{PP} = 300$ V, determine the Q point if (a) $R_k = 100$ ohms, (b) $R_k = 1$ K ohm.

PD-5. Given $R_L = 20$ K ohms and $V_{PP} = 300$ V, determine:

(a) the value of R_k to provide a bias of −2 V.
(b) the value of R_2 to provide a screen grid voltage of 150 V.
(c) the rms value of open circuit output voltage and voltage gain when a 1 V p-p input is applied.
(d) the rms value of output voltage and voltage gain with a 1 V p-p input and a 20 K-ohm load connected across the output.
(e) the dissipation of the tube under quiescent conditions and when a signal is applied as in (c).
(f) the small-signal voltage gain given $g_m = 3900$ μmhos and $r_p = 0.5$ M ohm, with $R_L = 20$ K ohms. Compare the result with (c).

PD-6. A cascaded pentode voltage amplifier utilizes two 6V6s in place of the 6AU6s in the circuit of Fig. D-6. Given: $C_{gp} = 0.3$, $C_i = C_{gk} = 10$, $C_o = C_{pk} = 11$ pF, $g_m = 3700$ μmhos, $r_p = 50$ K ohms, $R_g = 470$ K ohms and $R_L = 10$ K ohms. Assume identical stages. Calculate:

(a) the small-signal voltage gain of both stages at mid frequencies.
(b) the input capacitance to the second stage assuming no stray capacitance at all.
(c) the total shunting capacitance, C_T.
(d) the upper 3-dB frequency, f_2, for the first stage

(e) the voltage gain at this frequency, f_2.
(f) the gain-bandwidth product.
(g) the bandwidth if the gain is reduced by one-half (by a modification of R_L) from the value in (a).

PD-7. Repeat PD-6 assuming a stray capacitance of 1.7 pF is effectively connected from Grid 1 to the plate of the second tube.

PD-8. If a 6AU6 is to be operated in a cascaded RF voltage amplifier as in Fig. D-7, with a resonant frequency of 10.7 MHz and a bandwidth of 250 kHz, determine the effective value of shunting capacitance, C_T, if a voltage gain of 200 is required. $g_m = 3900$ μmhos.

PD-9. A 6AQ5 is used in a power amplifier circuit like that in Fig. D-10 with $V_{PP} = 275$ V, $R_k = 390$ ohms, $R_2 = 8.2$ K ohms, $R_g = 470$ K ohms, $C_k = 50\ \mu F$, $C_2 = 1\ \mu F$, $R_L = 16$ ohms, output transformer turns ratio = 20:1, and a primary dc resistance of 250 ohms. Analyze the amplifier as follows, using the characteristic curves of Fig. D-9:

(a) Draw the dc load line taking into account all resistances.
(b) Determine the operating point. Specify V_P, I_P, I_{G2}.
(c) Draw the ac load line.
(d) Determine the peak-to-peak output voltage and current at the transformer primary for a 10 V peak-to-peak input.
(e) Calculate the total harmonic distortion and the output power to R_L if the transformer is 75% efficient.
(f) Calculate the power gain in decibels.

PD-10. An RF amplifier with a voltage gain of 200 has a bandwidth of 250 kHz. The grid resistor at the input to the amplifier is 220 K ohms. What is the noise voltage developed in the grid resistor at room temperature, 20° C, and the noise voltage at the amplifier's output arising from this thermal noise?

E

APPENDIX

PHASE-SHIFT MEASUREMENT

The measurement of the phase shift may be done on an oscilloscope using Lissajous figures. See Fig. E-1.

Since v_s is normally the smaller voltage, it should be applied to the vertical input because it is the more sensitive.

The resulting oscilloscope display will give the phase shift if both v_s and v_1 give the same vertical and horizontal deflections when viewed separately on the oscilloscope.

Method:

(1) Apply v_s.

(2) Adjust vertical vernier to give some convenient height (ac input) for v_s.

(3) Center display in middle of 'scope vertically.

(4) Apply v_1 and temporarily disconnect v_s.

(5) Adjust horizontal vernier for same width in cm as v_s was adjusted for vertically.

(6) Center the display on the 'scope horizontally.

(7) Reapply v_s. The resulting display may be any of those in Fig. E-2.

Consider the ellipse shown in Fig. E-3. It can be shown that sine θ = sin C = OA/OB or AA'/BB', where θ is the angle of phase shift between v_s and v_1.

$$\therefore \theta = \text{arc sin} \frac{OA}{OB} .$$

For example, if OA = 0.7 cm and OB = 1 cm (or 1.4 cm and 2 cm), then

$$\theta = \text{arc sin} \frac{0.7}{1} \quad \text{arc sin } 0.7.$$

Looking in the sine table of trigonometric ratios for an angle whose sine is 0.7, we find 45°.

If the ellipse were sloping the other way, the angle determined for θ would be *subtracted* from 180° if it were known that the phase angle should be less than 180°, but it should be *added* to 180° if it is expected that the angle should be greater than 180°.

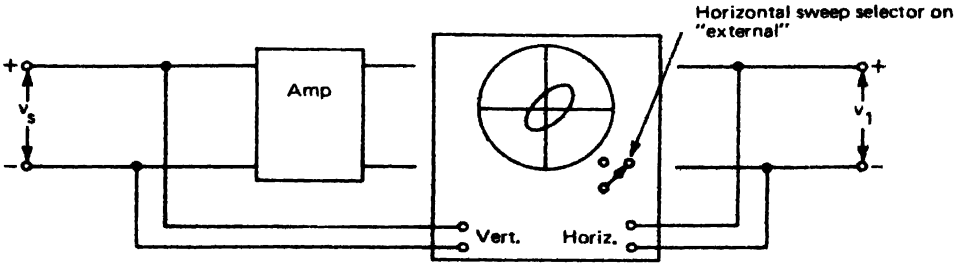


Fig. E-1. Oscilloscope setup for measurement of phase-shift angle.

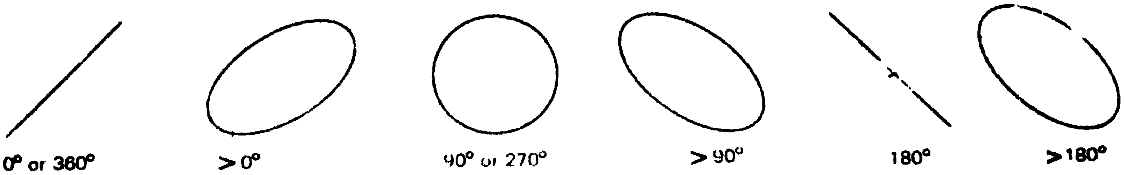

0° or 360°
> 0°
90° or 270°
> 90°
180°
> 180°

Fig. E-2. Some Lissajous figures corresponding to various phase-shift angles. Note: On some oscilloscopes the displays may slope the opposite way.

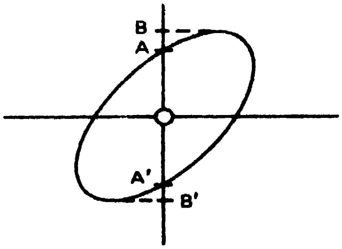


Fig. E-3. Showing the necessary measurements to be taken for phase-angle determination.

TABLE F 1

COMMON LOGARITHMS

x	0	1	2	3	4	5	6	7	8	9
1.0	.0000	.0043	.0086	.0128	.0170	.0212	.0253	.0294	.0334	.0374
1.1	.0414	.0453	.0492	.0531	.0569	.0607	.0645	.0682	.0719	.0755
1.2	.0792	.0828	.0864	.0899	.0934	.0969	.1004	.1038	.1072	.1106
1.3	.1139	.1173	.1206	.1239	.1271	.1303	.1335	.1367	.1399	.1430
1.4	.1461	.1492	.1523	.1553	.1584	.1614	.1644	.1673	.1703	.1732
1.5	.1761	.1790	.1818	.1847	.1875	.1903	.1931	.1959	.1987	.2014
1.6	.2041	.2068	.2095	.2122	.2148	.2175	.2201	.2227	.2253	.2279
1.7	.2304	.2330	.2355	.2380	.2405	.2430	.2455	.2480	.2504	.2529
1.8	.2553	.2577	.2601	.2625	.2648	.2672	.2695	.2718	.2742	.2765
1.9	.2788	.2810	.2833	.2856	.2878	.2900	.2923	.2945	.2967	.2989
2.0	.3010	.3032	.3054	.3075	.3096	.3118	.3139	.3160	.3181	.3201
2.1	.3222	.3243	.3263	.3284	.3304	.3324	.3345	.3365	.3385	.3404
2.2	.3424	.3444	.3464	.3483	.3502	.3522	.3541	.3560	.3579	.3598
2.3	.3617	.3636	.3655	.3674	.3692	.3711	.3729	.3747	.3766	.3784
2.4	.3802	.3820	.3838	.3856	.3874	.3892	.3909	.3927	.3945	.3962
2.5	.3979	.3997	.4014	.4031	.4048	.4065	.4082	.4099	.4116	.4133
2.6	.4150	.4166	.4183	.4200	.4216	.4232	.4249	.4265	.4281	.4298
2.7	.4314	.4330	.4346	.4362	.4378	.4393	.4409	.4425	.4440	.4456
2.8	.4472	.4487	.4502	.4518	.4533	.4548	.4564	.4579	.4594	.4609
2.9	.4624	.4639	.4654	.4669	.4683	.4698	.4713	.4728	.4742	.4757
3.0	.4771	.4786	.4800	.4814	.4829	.4843	.4857	.4871	.4886	.4900
3.1	.4914	.4928	.4942	.4955	.4969	.4983	.4997	.5011	.5024	.5038
3.2	.5051	.5065	.5079	.5092	.5105	.5119	.5132	.5145	.5159	.5172
3.3	.5185	.5198	.5211	.5224	.5237	.5250	.5263	.5276	.5289	.5302
3.4	.5315	.5328	.5340	.5353	.5366	.5378	.5391	.5403	.5416	.5428
3.5	.5441	.5453	.5465	.5478	.5490	.5502	.5514	.5527	.5539	.5551
3.6	.5563	.5575	.5587	.5599	.5611	.5623	.5635	.5647	.5658	.5670
3.7	.5682	.5694	.5705	.5717	.5729	.5740	.5752	.5763	.5775	.5786
3.8	.5798	.5809	.5821	.5832	.5843	.5855	.5866	.5877	.5888	.5899
3.9	.5911	.5922	.5933	.5944	.5955	.5966	.5977	.5988	.5999	.6010
4.0	.6021	.6031	.6042	.6053	.6064	.6075	.6085	.6096	.6107	.6117
4.1	.6128	.6138	.6149	.6160	.6170	.6180	.6191	.6201	.6212	.6222
4.2	.6232	.6243	.6253	.6263	.6274	.6284	.6294	.6304	.6314	.6325
4.3	.6335	.6345	.6355	.6365	.6375	.6385	.6395	.6405	.6415	.6425
4.4	.6435	.6444	.6454	.6464	.6474	.6484	.6493	.6503	.6513	.6522
4.5	.6532	.6542	.6551	.6561	.6571	.6580	.6590	.6599	.6609	.6618
4.6	.6628	.6637	.6646	.6656	.6665	.6675	.6684	.6693	.6702	.6712
4.7	.6721	.6730	.6739	.6749	.6758	.6767	.6776	.6785	.6794	.6803
4.8	.6812	.6821	.6830	.6839	.6848	.6857	.6866	.6875	.6884	.6893
4.9	.6902	.6911	.6920	.6928	.6937	.6946	.6955	.6964	.6972	.6981
5.0	.6990	.6998	.7007	.7016	.7024	.7033	.7042	.7050	.7059	.7067
5.1	.7076	.7084	.7093	.7101	.7110	.7118	.7126	.7135	.7143	.7152
5.2	.7160	.7168	.7177	.7185	.7193	.7202	.7210	.7218	.7226	.7235
5.3	.7243	.7251	.7259	.7267	.7275	.7284	.7292	.7300	.7308	.7316
5.4	.7324	.7332	.7340	.7348	.7356	.7364	.7372	.7380	.7388	.7396
x	0	1	2	3	4	5	6	7	8	9

TABLE I *(continued)*

x	0	1	2	3	4	5	6	7	8	9
5.5	.7404	.7412	.7419	.7427	.7435	.7443	.7451	.7459	.7466	.7474
5.6	.7482	.7490	.7497	.7505	.7513	.7520	.7528	.7536	.7543	.7551
5.7	.7559	.7566	.7574	.7582	.7589	.7597	.7604	.7612	.7619	.7627
5.8	.7634	.7642	.7649	.7657	.7664	.7672	.7679	.7686	.7694	.7701
5.9	.7709	.7716	.7723	.7731	.7738	.7745	.7752	.7760	.7767	.7774
6.0	.7782	.7789	.7796	.7803	.7810	.7818	.7825	.7832	.7839	.7846
6.1	.7853	.7860	.7868	.7875	.7882	.7889	.7896	.7903	.7910	.7917
6.2	.7924	.7931	.7938	.7945	.7952	.7959	.7966	.7973	.7980	.7987
6.3	.7993	.8000	.8007	.8014	.8021	.8028	.8035	.8041	.8048	.8055
6.4	.8062	.8069	.8075	.8082	.8089	.8096	.8102	.8109	.8116	.8122
6.5	.8129	.8136	.8142	.8149	.8156	.8162	.8169	.8176	.8182	.8189
6.6	.8195	.8202	.8209	.8215	.8222	.8228	.8235	.8241	.8248	.8254
6.7	.8261	.8267	.8274	.8280	.8287	.8293	.8299	.8306	.8312	.8319
6.8	.8325	.8331	.8338	.8344	.8351	.8357	.8363	.8370	.8376	.8382
6.9	.8388	.8395	.8401	.8407	.8414	.8420	.8426	.8432	.8439	.8445
7.0	.8451	.8457	.8463	.8470	.8476	.8482	.8488	.8494	.8500	.8506
7.1	.8513	.8519	.8525	.8531	.8537	.8543	.8549	.8555	.8561	.8567
7.2	.8573	.8579	.8585	.8591	.8597	.8603	.8609	.8615	.8621	.8627
7.3	.8633	.8639	.8645	.8651	.8657	.8663	.8669	.8675	.8681	.8686
7.4	.8692	.8698	.8704	.8710	.8716	.8722	.8727	.8733	.8739	.8745
7.5	.8751	.8756	.8762	.8768	.8774	.8779	.8785	.8791	.8797	.8802
7.6	.8808	.8814	.8820	.8825	.8831	.8837	.8842	.8848	.8854	.8859
7.7	.8865	.8871	.8876	.8882	.8887	.8893	.8899	.8904	.8910	.8915
7.8	.8921	.8927	.8932	.8938	.8943	.8949	.8954	.8960	.8965	.8971
7.9	.8976	.8982	.8987	.8993	.8998	.9004	.9009	.9015	.9020	.9025
8.0	.9031	.9036	.9042	.9047	.9053	.9058	.9063	.9069	.9074	.9079
8.1	.9085	.9090	.9096	.9101	.9106	.9112	.9117	.9122	.9128	.9133
8.2	.9138	.9143	.9149	.9154	.9159	.9165	.9170	.9175	.9180	.9186
8.3	.9191	.9196	.9201	.9206	.9212	.9217	.9222	.9227	.9232	.9238
8.4	.9243	.9248	.9253	.9258	.9263	.9269	.9274	.9279	.9284	.9289
8.5	.9294	.9299	.9304	.9309	.9315	.9320	.9325	.9330	.9335	.9340
8.6	.9345	.9350	.9355	.9360	.9365	.9370	.9375	.9380	.9385	.9390
8.7	.9395	.9400	.9405	.9410	.9415	.9420	.9425	.9430	.9435	.9440
8.8	.9445	.9450	.9455	.9460	.9465	.9469	.9474	.9479	.9484	.9489
8.9	.9494	.9499	.9504	.9509	.9513	.9518	.9523	.9528	.9533	.9538
9.0	.9542	.9547	.9552	.9557	.9562	.9566	.9571	.9576	.9581	.9586
9.1	.9590	.9595	.9600	.9605	.9609	.9614	.9619	.9624	.9628	.9633
9.2	.9638	.9643	.9647	.9652	.9657	.9661	.9666	.9671	.9675	.9680
9.3	.9685	.9689	.9694	.9699	.9703	.9708	.9713	.9717	.9722	.9727
9.4	.9731	.9736	.9741	.9745	.9750	.9754	.9759	.9763	.9768	.9773
9.5	.9777	.9782	.9786	.9791	.9795	.9800	.9805	.9809	.9814	.9818
9.6	.9823	.9827	.9832	.9836	.9841	.9845	.9850	.9854	.9859	.9863
9.7	.9868	.9872	.9877	.9881	.9886	.9890	.9894	.9899	.9903	.9908
9.8	.9912	.9917	.9921	.9926	.9930	.9934	.9939	.9943	.9948	.9952
9.9	.9956	.9961	.9965	.9969	.9974	.9978	.9983	.9987	.9991	.9996
x	0	1	2	3	4	5	6	7	8	9

TABLE II

EXPONENTIAL FUNCTIONS

x	e^x	e^{-x}	x	e^x	e^{-x}
0.00	1.0000	1.0000	1.5	4.4817	0.2231
0.01	1.0101	0.9901	1.6	4.9530	0.2019
0.02	1.0202	0.9802	1.7	5.4739	0.1827
0.03	1.0305	0.9705	1.8	6.0496	0.1653
0.04	1.0408	0.9608	1.9	6.6859	0.1496
0.05	1.0513	0.9512	2.0	7.3891	0.1353
0.06	1.0618	0.9418	2.1	8.1662	0.1225
0.07	1.0725	0.9324	2.2	9.0250	0.1108
0.08	1.0833	0.9331	2.3	9.9742	0.1003
0.09	1.0942	0.9139	2.4	11.023	0.0907
0.10	1.1052	0.9048	2.5	12.182	0.0821
0.11	1.1163	0.8958	2.6	13.464	0.0743
0.12	1.1275	0.8869	2.7	14.880	0.0672
0.13	1.1388	0.8781	2.8	16.445	0.0608
0.14	1.1503	0.8694	2.9	18.174	0.0550
0.15	1.1618	0.8607	3.0	20.086	0.0498
0.16	1.1735	0.8521	3.1	22.198	0.0450
0.17	1.1853	0.8437	3.2	24.533	0.0408
0.18	1.1972	0.8353	3.3	27.113	0.0369
0.19	1.2092	0.8270	3.4	29.964	0.0334
0.20	1.2214	0.8187	3.5	33.115	0.0302
0.21	1.2337	0.8106	3.6	36.598	0.0273
0.22	1.2461	0.8025	3.7	40.447	0.0247
0.23	1.2586	0.7945	3.8	44.701	0.0224
0.24	1.2712	0.7866	3.9	49.402	0.0202
0.25	1.2840	0.7788	4.0	54.598	0.0183
0.30	1.3499	0.7408	4.1	60.340	0.0166
0.35	1.4191	0.7047	4.2	66.686	0.0150
0.40	1.4918	0.6703	4.3	73.700	0.0136
0.45	1.5683	0.6376	4.4	81.451	0.0123
0.50	1.6487	0.6065	4.5	90.017	0.0111
0.55	1.7333	0.5769	4.6	99.484	0.0101
0.60	1.8221	0.5488	4.7	109.95	0.0091
0.65	1.9155	0.5220	4.8	121.51	0.0082
0.70	2.0138	0.4966	4.9	134.29	0.0074
0.75	2.1170	0.4724	5.0	148.41	0.0067
0.80	2.2255	0.4493	5.5	244.69	0.0041
0.85	2.3396	0.4274	6.0	403.43	0.0025
0.90	2.4596	0.4066	6.5	665.14	0.0015
0.95	2.5857	0.3867	7.0	096.6	0.0009
1.0	2.7183	0.3679	7.5	1808.0	0.0006
1.1	3.0042	0.3329	8.0	2981.0	0.0003
1.2	3.3201	0.3012	8.5	4914.8	0.0002
1.3	3.6693	0.2725	9.0	8103.1	0.0001
1.4	4.0552	0.2466	10.0	22026	0.00005

TABLE III

NATURAL LOGARITHMS OF NUMBERS

n	$\log_e n$	n	$\log_e n$	n	$\log_e n$
	*	4.5	1.5041	9.0	2.1972
0.1	7.6974	4.6	1.5261	9.1	2.2083
0.2	8.3906	4.7	1.5476	9.2	2.2192
0.3	8.7960	4.8	1.5686	9.3	2.2300
0.4	9.0837	4.9	1.5892	9.4	2.2407
0.5	9.3069	5.0	1.6094	9.5	2.2513
0.6	9.4892	5.1	1.6292	9.6	2.2618
0.7	9.6433	5.2	1.6487	9.7	2.2721
0.8	9.7769	5.3	1.6677	9.8	2.2824
0.9	9.8946	5.4	1.6864	9.9	2.2925
1.0	0.0000	5.5	1.7047	10	2.3026
1.1	0.0953	5.6	1.7228	11	2.3979
1.2	0.1823	5.7	1.7405	12	2.4849
1.3	0.2624	5.8	1.7579	13	2.5649
1.4	0.3365	5.9	1.7750	14	2.6391
1.5	0.4055	6.0	1.7918	15	2.7081
1.6	0.4700	6.1	1.8083	16	2.7726
1.7	0.5306	6.2	1.8245	17	2.8332
1.8	0.5878	6.3	1.8405	18	2.8904
1.9	0.6419	6.4	1.8563	19	2.9444
2.0	0.6931	6.5	1.8718	20	2.9957
2.1	0.7419	6.6	1.8871	25	3.2189
2.2	0.7885	6.7	1.9021	30	3.4012
2.3	0.8329	6.8	1.9169	35	3.5553
2.4	0.8755	6.9	1.9315	40	3.6889
2.5	0.9163	7.0	1.9459	45	3.8067
2.6	0.9555	7.1	1.9601	50	3.9120
2.7	0.9933	7.2	1.9741	55	4.0073
2.8	1.0296	7.3	1.9879	60	4.0943
2.9	1.0647	7.4	2.0015	65	4.1744
3.0	1.0986	7.5	2.0149	70	4.2485
3.1	1.1314	7.6	2.0281	75	4.3175
3.2	1.1632	7.7	2.0412	80	4.3820
3.3	1.1939	7.8	2.0541	85	4.4427
3.4	1.2238	7.9	2.0669	90	4.4998
3.5	1.2528	8.0	2.0794	100	4.6052
3.6	1.2809	8.1	2.0919	110	4.7005
3.7	1.3083	8.2	2.1041	120	4.7875
3.8	1.3350	8.3	2.1163	130	4.8676
3.9	1.3610	8.4	2.1282	140	4.9416
4.0	1.3863	8.5	2.1401	150	5.0106
4.1	1.4110	8.6	2.1518	160	5.0752
4.2	1.4351	8.7	2.1633	170	5.1358
4.3	1.4586	8.8	2.1748	180	5.1930
4.4	1.4816	8.9	2.1861	190	5.2470

Subtract 10 for $n < 1$. Thus $\log_e 0.1 = 7.6974 - 10$ -2.3026

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0°	.0000	0017	0035	0052	0070	.0087	0105	0122	0140	.0157
1°	0175	0192	0209	0227	0244	0262	0279	0297	0314	0332
2°	0349	0366	0384	0401	0419	0436	0454	0471	0488	0506
3°	0523	0541	0558	0576	0593	0610	0628	0645	0663	0680
4°	0698	0715	0732	0750	0767	0785	0802	0819	0837	0854
5°	0872	0889	0906	0924	0941	0958	0976	0993	1011	1028
6°	1045	1063	1080	1097	1115	1132	1149	1167	1184	1201
7°	1219	1236	1253	1271	1288	1305	1323	1340	1357	1374
8°	1392	1409	1426	1444	1461	1478	1495	1513	1530	1547
9°	1564	1582	1599	1616	1633	1650	1668	1685	1702	1719
10°	1736	1754	1771	1788	1805	1822	1840	1857	1874	1891
11°	1908	1925	1942	1959	1977	1994	2011	2028	2045	2062
12°	2079	2096	2113	2130	2147	2164	2181	2198	2215	2233
13°	2250	2267	2284	2300	2317	2334	2351	2368	2385	2402
14°	2419	2436	2453	2470	2487	2504	2521	2538	2554	2571
15°	2588	2605	2622	2639	2656	2672	2689	2706	2723	2740
16°	2756	2773	2790	2807	2823	2840	2857	2874	2890	2907
17°	2924	2940	2957	2974	2990	3007	3024	3040	3057	3074
18°	3090	3107	3123	3140	3156	3173	3190	3206	3223	3239
19°	3256	3272	3289	3305	3322	3338	3355	3371	3387	3404
20°	3420	3437	3453	3469	3486	3502	3518	3535	3551	3567
21°	3584	3600	3616	3633	3649	3665	3681	3697	3714	3730
22°	3746	3762	3778	3795	3811	3827	3843	3859	3875	3891
23°	3907	3923	3939	3955	3971	3987	4003	4019	4035	4051
24°	4067	4083	4099	4115	4131	4147	4163	4179	4195	4210
25°	4226	4242	4258	4274	4289	4305	4321	4337	4352	4368
26°	4384	4399	4415	4431	4446	4462	4478	4493	4509	4524
27°	4540	4555	4571	4586	4602	4617	4633	4648	4664	4679
28°	4695	4710	4726	4741	4756	4772	4787	4802	4818	4833
29°	4848	4863	4879	4894	4909	4924	4939	4955	4970	4985
30°	5000	5015	5030	5045	5060	5075	5090	5105	5120	5135
31°	5150	5165	5180	5195	5210	5225	5240	5255	5270	5284
32°	5299	5314	5329	5344	5358	5373	5388	5402	5417	5432
33°	5446	5461	5476	5490	5505	5519	5534	5548	5563	5577
34°	5592	5606	5621	5635	5650	5664	5678	5693	5707	5721
35°	5736	5750	5764	5779	5793	5807	5821	5835	5850	5864
36°	5878	5892	5906	5920	5934	5948	5962	5976	5990	6004
37°	6018	6032	6046	6060	6074	6088	6101	6115	6129	6143
38°	6157	6170	6184	6198	6211	6225	6239	6252	6266	6280
39°	6293	6307	6320	6334	6347	6361	6374	6388	6401	6414
40°	6428	6441	6455	6468	6481	6494	6508	6521	6534	6547
41°	6561	6574	6587	6600	6613	6626	6639	6652	6665	6678
42°	6691	6704	6717	6730	6743	6756	6769	6782	6794	6807
43°	6820	6833	6845	6858	6871	6884	6896	6909	6921	6934
44°	6947	6959	6972	6984	6997	7009	7022	7034	7046	7059

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
45°	.7071	.7083	.7096	.7108	.7120	.7133	.7145	.7157	.7169	.7181
46°	.7193	.7206	.7218	.7230	.7242	.7254	.7266	.7278	.7290	.7302
47°	.7314	.7325	.7337	.7349	.7361	.7373	.7385	.7396	.7408	.7420
48°	.7431	.7443	.7455	.7466	.7478	.7490	.7501	.7513	.7524	.7536
49°	.7547	.7559	.7570	.7581	.7593	.7604	.7615	.7627	.7638	.7649
50°	.7660	.7672	.7683	.7694	.7705	.7716	.7727	.7738	.7749	.7760
51°	.7771	.7782	.7793	.7804	.7815	.7826	7837	.7848	.7859	.7869
52°	.7880	.7891	.7902	.7912	.7923	.7934	7944	.7955	.7965	.7976
53°	.7986	.7997	.8007	.8018	.8028	.8039	8049	.8059	.8070	.8080
54°	.8090	.8100	.8111	.8121	.8131	.8141	8151	.8161	.8171	.8181
55°	.8192	.8202	.8211	.8221	.8231	.8241	.8251	.8261	.8271	.8281
56°	.8290	.8300	.8310	.8320	.8329	.8339	.8348	.8358	.8368	.8377
57°	.8387	.8396	.8406	.8415	.8425	.8434	8443	8453	.8462	.8471
58°	.8480	.8490	.8499	.8508	.8517	.8526	.8536	.8545	.8554	.8563
59°	.8572	.8581	.8590	.8599	.8607	.8616	.8625	.8634	.8643	.8652
60°	.8660	.8669	.8678	.8686	.8695	.8704	.8712	.8721	.8729	.8738
61°	.8746	.8755	.8763	.8771	.8780	.8788	.8796	.8805	.8813	.8821
62°	.8829	.8838	.8846	8854	.8862	8870	.8878	.8886	.8894	.8902
63°	.8910	.8918	.8926	.8934	.8942	8949	.8957	8965	.8973	.8980
64°	.8988	.8996	.9003	.9011	.9018	.9026	.9033	.9041	.9048	.9056
65°	.9063	.9070	.9078	.9085	.9092	.9100	.9107	.9114	.9121	.9128
66°	.9135	.9143	.9150	.9157	.9164	9171	9178	.9184	.9191	.9198
67°	.9205	.9212	.9219	.9225	.9232	.9239	9245	.9252	.9259	.9265
68°	.9272	.9278	.9285	.9291	.9298	9304	.9311	9317	.9323	.9330
69°	.9336	.9342	.9348	.9354	.9361	.9367	.9373	9379	.9385	.9391
70°	.9397	.9403	.9409	.9415	.9421	.9426	9432	.9438	.9444	.9449
71°	.9455	.9461	.9466	.9472	.9478	.9483	.9489	9494	.9500	.9505
72°	.9511	.9516	9521	.9527	.9532	.9537	.9542	.9548	.9553	.9558
73°	.9563	.9568	.9573	.9578	.9583	.9588	.9593	9598	.9603	.9608
74°	.9613	.9617	.9622	.9627	.9632	.9636	.9641	.9646	.9650	.9655
75°	.9659	9664	.9668	.9673	.9677	.9681	.9686	.9690	.9694	.9699
76°	.9703	.9707	.9711	.9715	.9720	.9724	.9728	.9732	.9736	9740
77°	.9744	.9748	.9751	.9755	.9759	9763	.9767	.9770	.9774	.9778
78°	.9781	.9785	9789	.9792	.9796	.9799	.9803	9806	9810	.9813
79°	.9816	.9820	.9823	.9826	.9829	.9833	.9836	.9839	.9842	.9845
80°	.9848	9851	.9854	9857	.9860	9863	.9866	.9869	.9871	.9874
81°	.9877	9880	.9882	9885	.9888	.9890	.9893	.9895	.9898	.9900
82°	.9903	.9905	.9907	.9910	.9912	9914	9917	9919	.9921	.9923
83°	.9925	.9928	9930	.9932	.9934	.9936	.9938	.9940	.9942	.9943
84°	.9945	.9947	.9949	.9951	.9952	.9954	.9956	.9957	.9959	.9960
85°	.9962	.9963	9965	9966	.9968	.9969	.9971	.9972	.9973	.9974
86°	.9976	.9977	.9978	.9979	.9980	9981	.9982	.9983	.9984	.9985
87°	.9986	.9987	9988	.9989	.9990	.9990	.9991	.9992	.9993	.9993
88°	.9994	.9995	.9995	.9996	.9996	.9997	.9997	.9997	.9998	.9998
89°	.9998	.9999	.9999	.9999	.9999	1.000	1.000	1.000	1.000	1.000

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0°	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	.9999	.9999	.9999	.9999
1°	.9998	.9998	.9998	.9997	.9997	.9997	.9996	.9996	.9995	.9995
2°	.9994	.9993	.9993	.9992	.9991	.9990	.9990	.9989	.9988	.9987
3°	.9986	.9985	.9984	.9983	.9982	.9981	.9980	.9979	.9978	.9977
4°	.9976	.9974	.9973	.9972	.9971	.9969	.9968	.9966	.9965	.9963
5°	.9962	.9960	.9959	.9957	.9956	.9954	.9952	.9951	.9949	.9947
6°	.9945	.9943	.9942	.9940	.9938	.9936	.9934	.9932	.9930	.9928
7°	.9925	.9923	.9921	.9919	.9917	.9914	.9912	.9910	.9907	.9905
8°	.9903	.9900	.9898	.9895	.9893	.9890	.9888	.9885	.9882	.9880
9°	.9877	.9874	.9871	.9869	.9866	.9863	.9860	.9857	.9854	.9851
10°	.9848	.9845	.9842	.9839	.9836	.9833	.9829	.9826	.9823	.9820
11°	.9816	.9813	.9810	.9806	.9803	9799	.9796	.9792	.9789	.9785
12°	.9781	.9778	.9774	.9770	.9767	9763	.9759	.9755	.9751	.9748
13°	.9744	.9740	.9736	.9732	.9728	.9724	.9720	9715	.9711	.9707
14°	.9703	.9699	.9694	.9690	.9686	.9681	.9677	.9673	.9668	.9664
15°	.9659	.9655	.9650	.9646	.9641	.9636	.9632	.9627	.9622	.9617
16°	.9613	.9608	.9603	.9598	.9593	.9588	.9583	.9578	.9573	.9568
17°	.9563	.9558	.9553	.9548	.9542	.9537	.9532	.9527	.9521	.9516
18°	.9511	.9505	.9500	.9494	.9489	.9483	.9478	.9472	.9466	.9461
19°	.9455	.9449	.9444	.9438	.9432	.9426	.9421	.9415	.9409	.9403
20°	.9397	.9391	.9385	.9379	.9373	.9367	.9361	.9354	.9348	.9342
21°	.9336	.9330	.9323	.9317	.9311	.9304	.9298	9291	.9285	.9278
22°	.9272	.9265	.9259	.9252	.9245	.9239	.9232	.9225	.9219	.9212
23°	.9205	.9198	.9191	.9184	.9178	.9171	.9164	.9157	.9150	.9143
24°	.9135	.9128	.9121	.9114	.9107	.9100	.9092	.9085	.9078	.9070
25°	.9063	.9056	.9048	.9041	.9033	.9026	.9018	.9011	.9003	.8996
26°	.8988	.8980	.8973	.8965	.8957	.8949	.8942	.8934	.8926	.8918
27°	.8910	.8902	.8894	.8886	.8878	.8870	.8862	.8854	.8846	.8838
28°	.8829	.8821	.8813	.8805	.8796	.8788	.8780	.8771	.8763	.8755
29°	.8746	.8738	.8729	.8721	.8712	.8704	.8695	.8686	.8678	.8669
30°	.8660	.8652	.8643	.8634	.8625	.8616	.8607	.8599	.8590	.8581
31°	.8572	.8563	.8554	.8545	.8536	.8526	.8517	.8508	.8499	.8490
32°	.8480	.8471	.8462	.8453	.8443	.8434	.8425	.8415	.8406	.8396
33°	.8387	.8377	.8368	.8358	.8348	.8339	.8329	.8320	.8310	.8300
34°	.8290	.8281	.8271	.8261	.8251	.8241	.8231	.8221	.8211	.8202
35°	.8192	.8181	.8171	.8161	.8151	.8141	.8131	.8121	.8111	.8100
36°	.8090	.8080	.8070	.8059	.8049	.8039	.8028	.8018	.8007	.7997
37°	.7986	.7976	.7965	.7955	.7944	.7934	.7923	.7912	.7902	.7891
38°	.7880	.7869	.7859	.7848	.7837	.7826	.7815	.7804	.7793	.7782
39°	.7771	.7760	.7749	.7738	.7727	.7716	.7705	.7694	.7683	.7672
40°	.7660	.7649	.7638	.7627	.7615	.7604	.7593	.7581	.7570	.7559
41°	.7547	.7536	.7524	.7513	.7501	.7490	.7478	.7466	.7455	.7443
42°	.7431	.7420	.7408	.7396	.7385	.7373	.7361	.7349	.7337	.7325
43°	.7314	.7302	.7290	.7278	.7266	.7254	.7242	.7230	.7218	.7206
44°	.7193	.7181	.7169	.7157	.7145	.7133	.7120	.7108	.7096	.7083

Natural Trigonometric Functions: Cosine

Angles from 45°.0 to 89°.9

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
45°	.7071	7059	7046	.7034	7022	.7009	6997	.6984	.6972	.6959
46°	6947	6934	.6921	.6909	.6896	6884	6871	.6858	.6845	6833
47°	6820	6807	.6794	.6782	.6769	6756	.6743	6730	.6717	6704
48°	.6691	6678	.6665	6652	.6639	.6626	6613	.6600	.6587	.6574
49°	6561	6547	6534	.6521	.6508	.6494	6481	6468	.6455	.6441
50°	.6428	6414	.6401	6388	6374	6361	.6347	.6334	.6320	.6307
51°	6293	6280	.6266	.6252	6239	.6225	6211	6198	.6184	6170
52°	6157	.6143	.6129	.6115	6101	6088	.6074	6060	6046	6032
53°	.6018	6004	.5990	5976	5962	5948	5934	.5920	.5906	.5892
54°	5878	5864	5850	5835	5821	5807	5793	5779	.5764	.5750
55°	.5736	5721	5707	.5693	5678	5664	.5650	.5635	.5621	.5606
56°	5592	5577	5563	.5548	5534	5519	.5505	.5490	.5476	.5461
57°	5446	5432	.5417	.5402	.5388	.5373	.5358	5344	5329	.5314
58°	.5299	5284	.5270	.5255	.5240	5225	.5210	.5195	.5180	.5165
59°	.5150	5135	.5120	.5105	5090	5075	5060	5045	.5030	.5015
60°	.5000	.4985	.4970	.4955	.4939	4924	4909	.4894	.4879	.4863
61°	.4848	.4833	4818	.4802	4787	4772	4756	.4741	.4726	.4710
62°	.4695	.4679	.4664	.4648	4633	4617	.4602	.4586	4571	.4555
63°	.4540	.4524	4509	.4493	4478	4462	.4446	.4431	.4415	.4399
64°	4384	4368	.4352	4337	4321	4305	.4289	4274	.4258	.4242
65°	.4226	.4210	.4195	4179	4163	4147	.4131	.4115	.4099	.4083
66°	.4067	4051	4035	4019	4003	3987	3971	3955	.3939	3923
67°	.3907	3891	3875	.3859	.3843	3827	.3811	.3795	.3778	.3762
68°	3746	3730	.3714	3697	3681	.3665	3649	.3633	.3616	3600
69°	3584	.3567	.3551	3535	3518	3502	3486	.3469	.3453	.3437
70°	.3420	.3404	.3387	.3371	.3355	.3338	.3322	.3305	.3289	3272
71°	.3256	.3239	3223	.3206	.3190	3173	3156	.3140	3123	.3107
72°	3090	3074	.3057	3040	3024	.3007	.2990	.2974	2957	.2940
73°	.2924	.2907	.2890	.2874	.2857	2840	2823	.2807	2790	2773
74°	.2756	.2740	.2723	2706	2689	2672	.2656	2639	.2622	2605
75°	.2588	.2571	.2554	.2538	2521	2504	.2487	.2470	.2453	.2436
76°	.2419	2402	.2385	.2368	2351	.2334	2317	.2300	2284	2267
77°	.2250	2233	.2215	.2198	2181	.2164	.2147	.2130	.2113	2096
78°	.2079	2062	.2045	2028	2011	1994	.1977	.1959	.1942	1925
79°	.1908	.1891	.1874	.1857	1840	.1822	.1805	.1788	.1771	.1754
80°	.1736	.1719	.1702	.1685	1668	1650	.1633	.1616	.1599	.1582
81°	.1564	.1547	1530	.1513	1495	.1478	.1461	.1444	.1426	.1409
82°	.1392	.1374	.1357	.1340	.1323	.1305	.1288	.1271	.1253	.1236
83°	.1219	.1201	.1184	.1167	.1149	.1132	.1115	.1097	.1080	.1063
84°	.1045	.1028	.1011	.0993	0976	.0958	.0941	.0924	.0906	.0889
85°	.0872	.0854	.0837	.0819	.0802	.0785	.0767	.0750	.0732	.0715
86°	.0698	.0680	.0663	0645	0628	0610	.0593	.0576	.0558	.0541
87°	.0523	.0506	.0488	.0471	0454	0436	.0419	.0401	.0384	.0366
88°	.0349	0332	.0314	.0297	.0279	.0262	.0244	.0227	.0209	.0192
89°	.0175	.0157	.0140	.0122	.0105	.0087	.0070	.0052	.0035	.0017

Natural Trigonometric Functions: Tangent

Angles from 0°.0 to 44°.9

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
0°	.0000	.0017	.0035	.0052	.0070	.0087	.0105	.0122	.0140	.0157
1°	.0175	.0192	.0209	.0227	.0244	.0262	.0279	.0297	.0314	.0332
2°	.0349	.0367	.0384	.0402	.0419	.0437	.0454	.0472	.0489	.0507
3°	.0524	.0542	.0559	.0577	.0594	.0612	.0629	.0647	.0664	.0682
4°	.0699	.0717	.0734	.0752	.0769	.0787	.0805	.0822	.0840	.0857
5°	.0875	.0892	.0910	.0928	.0945	.0963	.0981	.0998	.1016	.1033
6°	.1051	.1069	.1086	.1104	.1122	.1139	.1157	.1175	.1192	.1210
7°	.1228	.1246	.1263	.1281	.1299	.1317	.1334	.1352	.1370	.1388
8°	.1405	.1423	.1441	.1459	.1477	.1495	.1512	.1530	.1548	.1566
9°	.1584	.1602	.1620	.1638	.1655	.1673	.1691	.1709	.1727	.1745
10°	.1763	.1781	.1799	.1817	.1835	.1853	.1871	.1890	.1908	.1926
11°	.1944	.1962	.1980	.1998	.2016	.2035	.2053	.2071	.2089	.2107
12°	.2126	.2144	.2162	.2180	.2199	.2217	.2235	.2254	.2272	.2290
13°	.2309	.2327	.2345	.2364	.2382	.2401	.2419	.2438	.2456	.2475
14°	.2493	.2512	.2530	.2549	.2568	.2586	.2605	.2623	.2642	.2661
15°	.2679	.2698	.2717	.2736	.2754	.2773	.2792	.2811	.2830	.2849
16°	.2867	.2886	.2905	.2924	.2943	.2962	.2981	.3000	.3019	.3038
17°	.3057	.3076	.3096	.3115	.3134	.3153	.3172	.3191	.3211	.3230
18°	.3249	.3269	.3288	.3307	.3327	.3346	.3365	.3385	.3404	.3424
19°	.3443	.3463	.3482	.3502	.3522	.3541	.3561	.3581	.3600	.3620
20°	.3640	.3659	.3679	.3699	.3719	.3739	.3759	.3779	.3799	.3819
21°	.3839	.3859	.3879	.3899	.3919	.3939	.3959	.3979	.4000	.4020
22°	.4040	.4061	.4081	.4101	.4122	.4142	.4163	.4183	.4204	.4224
23°	.4245	.4265	.4286	.4307	.4327	.4348	.4369	.4390	.4411	.4431
24°	.4452	.4473	.4494	.4515	.4536	.4557	.4578	.4599	.4621	.4642
25°	.4663	.4684	.4706	.4727	.4748	.4770	.4791	.4813	.4834	.4856
26°	.4877	.4899	.4921	.4942	.4964	.4986	.5008	.5029	.5051	.5073
27°	.5095	.5117	.5139	.5161	.5184	.5206	.5228	.5250	.5272	.5295
28°	.5317	.5340	.5362	.5384	.5407	.5430	.5452	.5475	.5498	.5520
29°	.5543	.5566	.5589	.5612	.5635	.5658	.5681	.5704	.5727	.5750
30°	.5774	.5797	.5820	.5844	.5867	.5890	.5914	.5938	.5961	.5985
31°	.6009	.6032	.6056	.6080	.6104	.6128	.6152	.6176	.6200	.6224
32°	.6249	.6273	.6297	.6322	.6346	.6371	.6395	.6420	.6445	.6469
33°	.6494	.6519	.6544	.6569	.6594	.6619	.6644	.6669	.6694	.6720
34°	.6745	.6771	.6796	.6822	.6847	.6873	.6899	.6924	.6950	.6976
35°	.7002	.7028	.7054	.7080	.7107	.7133	.7159	.7186	.7212	.7239
36°	.7265	.7292	.7319	.7346	.7373	.7400	.7427	.7454	.7481	.7508
37°	.7536	.7563	.7590	.7618	.7646	.7673	.7701	.7729	.7757	.7785
38°	.7813	.7841	.7869	.7898	.7926	.7954	.7983	.8012	.8040	.8069
39°	.8098	.8127	.8156	.8185	.8214	.8243	.8273	.8302	.8332	.8361
40°	.8391	.8421	.8451	.8481	.8511	.8541	.8571	.8601	.8632	.8662
41°	.8693	.8724	.8754	.8785	.8816	.8847	.8878	.8910	.8941	.8972
42°	.9004	.9036	.9067	.9099	.9131	.9163	.9195	.9228	.9260	.9293
43°	.9325	.9358	.9391	.9424	.9457	.9490	.9523	.9556	.9590	.9623
44°	.9657	.9691	.9725	.9759	.9793	.9827	.9861	.9896	.9930	.9965

	.0	.1	.2	.3	.4	.5	.6	.7	.8	.9
45°	1.000	1.003	1.007	1.011	1.014	1.018	1.021	1.025	1.028	1.032
46°	1.036	1.039	1.043	1.046	1.050	1.054	1.057	1.061	1.065	1.069
47°	1.072	1.076	1.080	1.084	1.087	1.091	1.095	1.099	1.103	1.107
48°	1.111	1.115	1.118	1.122	1.126	1.130	1.134	1.138	1.142	1.146
49°	1.150	1.154	1.159	1.163	1.167	1.171	1.175	1.179	1.183	1.188
50°	1.192	1.196	1.200	1.205	1.209	1.213	1.217	1.222	1.226	1.230
51°	1.235	1.239	1.244	1.248	1.253	1.257	1.262	1.266	1.271	1.275
52°	1.280	1.285	1.289	1.294	1.299	1.303	1.308	1.313	1.317	1.322
53°	1.327	1.332	1.337	1.342	1.347	1.351	1.356	1.361	1.366	1.371
54°	1.376	1.381	1.387	1.392	1.397	1.402	1.407	1.412	1.418	1.423
55°	1.428	1.433	1.439	1.444	1.450	1.455	1.460	1.466	1.471	1.477
56°	1.483	1.488	1.494	1.499	1.505	1.511	1.517	1.522	1.528	1.534
57°	1.540	1.546	1.552	1.558	1.564	1.570	1.576	1.582	1.588	1.594
58°	1.600	1.607	1.613	1.619	1.625	1.632	1.638	1.645	1.651	1.658
59°	1.664	1.671	1.678	1.684	1.691	1.698	1.704	1.711	1.718	1.725
60°	1.732	1.739	1.746	1.753	1.760	1.767	1.775	1.782	1.789	1.797
61°	1.804	1.811	1.819	1.827	1.834	1.842	1.849	1.857	1.865	1.873
62°	1.881	1.889	1.897	1.905	1.913	1.921	1.929	1.937	1.946	1.954
63°	1.963	1.971	1.980	1.988	1.997	2.006	2.014	2.023	2.032	2.041
64°	2.050	2.059	2.069	2.078	2.087	2.097	2.106	2.116	2.125	2.135
65°	2.145	2.154	2.164	2.174	2.184	2.194	2.204	2.215	2.225	2.236
66°	2.246	2.257	2.267	2.278	2.289	2.300	2.311	2.322	2.333	2.344
67°	2.356	2.367	2.379	2.391	2.402	2.414	2.426	2.438	2.450	2.463
68°	2.475	2.488	2.500	2.513	2.526	2.539	2.552	2.565	2.578	2.592
69°	2.605	2.619	2.633	2.646	2.660	2.675	2.689	2.703	2.718	2.733
70°	2.747	2.762	2.778	2.793	2.808	2.824	2.840	2.856	2.872	2.888
71°	2.904	2.921	2.937	2.954	2.971	2.989	3.006	3.024	3.042	3.060
72°	3.078	3.096	3.115	3.133	3.152	3.172	3.191	3.211	3.230	3.251
73°	3.271	3.291	3.312	3.333	3.354	3.376	3.398	3.420	3.442	3.465
74°	3.487	3.511	3.534	3.558	3.582	3.606	3.630	3.655	3.681	3.706
75°	3.732	3.758	3.785	3.812	3.839	3.867	3.895	3.923	3.952	3.981
76°	4.011	4.041	4.071	4.102	4.134	4.165	4.198	4.230	4.264	4.297
77°	4.331	4.366	4.402	4.437	4.474	4.511	4.548	4.586	4.625	4.665
78°	4.705	4.745	4.787	4.829	4.872	4.915	4.959	5.005	5.050	5.097
79°	5.145	5.193	5.242	5.292	5.343	5.396	5.449	5.503	5.558	5.614
80°	5.671	5.730	5.789	5.850	5.912	5.976	6.041	6.107	6.174	6.243
81°	6.314	6.386	6.460	6.535	6.612	6.691	6.772	6.855	6.940	7.026
82°	7.115	7.207	7.300	7.396	7.495	7.596	7.700	7.806	7.916	8.028
83°	8.144	8.264	8.386	8.513	8.643	8.777	8.915	9.058	9.205	9.357
84°	9.514	9.677	9.845	10.02	10.20	10.39	10.58	10.78	10.99	11.20
85°	11.43	11.66	11.91	12.16	12.43	12.71	13.00	13.30	13.62	13.95
86°	14.30	14.67	15.06	15.46	15.89	16.35	16.83	17.34	17.89	18.46
87°	19.08	19.74	20.45	21.20	22.02	22.90	23.86	24.90	26.03	27.27
88°	28.64	30.14	31.82	33.69	35.80	38.19	40.92	44.07	47.74	52.08
89°	57.29	63.66	71.62	81.85	95.49	114.6	143.2	191.0	286.5	573.0

ANSWERS TO ODD-NUMBERED PROBLEMS

(For answers involving graphs and waveforms refer to the Instructor's Manual)

CHAPTER 1

1. 6; 7
3. (a) $I_p = KV_p{}^{3/2}$; I_p in mA, V_p in volts
   (b) 97.6 mA
   (c) 10 V
5. 16 V
7. 0.154
9. (a) $R_p = 385$ ohms
   (b) $V_o = 25$ V; $r_p = 283$ ohms
15. 150 V
17. (a) 1 mA
   (b) 1.56 mA
   (c) 1.56 mA

CHAPTER 2

1. 8.4 mA
5. (a) $I_{rms} \approx \sqrt{\left(\frac{I_m}{\pi}\right)^2 + \left(\frac{I_m}{2\sqrt{2}}\right)^2} = 0.477I_m$
   (b) $I_r \approx \frac{\sqrt{\left(\frac{I_m}{\pi}\right)^2 + \left(\frac{I_m}{2\sqrt{2}}\right)^2}}{\left(\frac{2\,I_m}{3\pi\sqrt{2}}\right)^2} = 0.499I_m$
7. (a) 5.66 V
   (b) 5.26 V
   (c) 7.6%
   (d) 4 ohms
9. (a) 20 V
   (b) 25 V
   (c) 40 V
11. (a) 1232 W
   (b) 616 W
13. (a) 348 V A.
   (b) 144 V.A.
   (c) 123 V.A.
   (d) 105 V.A.
15. (a) 270 V
   (b) 270 mA
   (c) 73 W
   (d) 105 V.A.
   (e) 113 W
   (f) 130 V, 120 Hz
   (g) 150 V
   (h) 848 V
   (i) 85 W
   (j) 540 V
   (k) 848 V
17. (a) 10 V
   (b) 8.525 V

CHAPTER 3

1. 1.25 H
3. 45 V
5. (b) 11.8%
7. (a) 2.41%
   (b) 8.55 V
9. (a) 86 ohms
   (b) approximately 28% (18% more accurately)
11. 0.91 A
13. $r = 0.25$; $I_p = 1.02$ A
15. (a) $r = 0.18$; $I_p = 0.766$ A

(b) 0.9 A
17. 0.91 H; 91 μF
19. (a) L=1.66 H; C=100 μF
(b) 1500 ohms
(c) 32.2 V, C.T.
(d) 46 V
(e) 0.256 A
21. 1.46×10^{-4}
23. (a) 1.32 H
(b) 1000 ohms
(c) 3 V

CHAPTER 4

1. (a) −10 V
(b) +10 V
(c) −20 V
(d) 565 V
3. (a) 283 V
(b) C_1 to 141.4 V; All others to 283 V
5. (a) 3250 V
(b) C_1 to 1630 V; C_2 to 3260 V
(c) 0.732%
7. (a) 70.7 V
(b) 1.1%
(c) 60 Hz; 120 Hz
(d) 200 V
(e) 100 V
(f) 0.21 A

CHAPTER 5

1. (a) 4 mA; 30 mA; 219 mA; 592 mA; 1.6 A
3. 91.4 μA
5. (a) 0.493 V
(b) 0.652 V
7. (a) 0.144 μA
(b) 3.17 mA
(c) 8.65 mA
9. 3.0%

CHAPTER 6

1. (a) 15 ohms
(b) 30 mV
(c) 10 mA
3. (a) 25 V
(b) 17.6 V
5. Between 10 mA and 40 mA.
7. (a) −2.5%/°C
(b) −1%/°C
9. (a) 88°C
(b) 225 sec.

CHAPTER 7

1. (a) $A_i = 100$; $A_v = 0.9$; $A_p = 90$
(b) 50 K ohms
3. (a) 6 V
(b) 5.4 V
5. (a) 163 μW @ 200 ohms
(b) 2 V @ 1 M ohm
7. (a) 5 V p-p
(b) 2.86 V p-p
(c) 10 V p-p
9. 180 hr
11. (a) 20 mA
(b) 1 mA; 5%
(c) 47 mA, 7 mA
(d) 0.2 W
(e) 0.2 W
13. (a) 5.7%
(b) (Eq. 7-26) 8.6 mW; (Eq. 7-27) 7.8 mW; (Eq. 7-28) 8.6 mW

CHAPTER 8

1. (a) 10:1

  (b) 1.26:1
3. 0.4 mW
5. (a) 63 dB
  (b) 34.8 dB
  (c) 63 dB
7. 10^4:1
9. (a) 22.2 dB
  (b) 30 dB
11. (a) 0.079 W
  (b) 0.25 W
  (c) 0.125 W
13. (a) 40 dB
  (b) 18 V; 8.1 ohms
  (c) 39.9 dB
15. (a) 140 dB
  (b) 1×10^{-11} W
17. (a) $0.062 \angle{-51^\circ}$
  (b) $0.312 \angle{-64.2^\circ}$; dB gain = 14.04 dB
  (c) $0.0127 \angle{-80^\circ}$; dB gain = −13.8 dB
19. 30 dB

CHAPTER 9

1. (a) 125
  (b) 172
  (c) 190
3. (a) 97
  (b) 155
  (c) 165
5. (a) 1 mA
  (b) For silicon, I_{CEO} is much smaller and difference would be negligible.
7. 0.995
9. 28.5 @ 10 mA; 23.57 @ 100 mA; 10.83 @ 600 mA
11. $h_{FE\,avg} = 20.97$; $h_{FB} = 0.955$; $h_{fe\,avg} = 16.39$; $h_{fb} = 0.942$
13. (a) 10.714 mA
  (b) 32.8 mA
  (c) 49 mA

CHAPTER 10

1. (b) 2.8 K ohms
  (c) 15 V, 3 mA
  (d) 0.54 V
  (e) 180 ohms
  (f) 20 V p-p
  (g) 0.0625 V p-p
  (h) 320
  (i) 3.95 mA p-p
  (j) 0.988
  (k) 15.6 ohms
  (l) 316
  (m) Slope = 3 mA/10 V
  (n) 13.2 V p-p
  (o) 211
  (p) 0.988
  (q) 23.2 dB
3. (a) 15.6 ohms
  (b) 4.76 K ohms
  (c) 0.982; No
  (d) (i) 320
    (ii) 194
    (iii) 76
5. (a) 0.0313 mW
  (b) 9.88 mW
  (c) 316; 25 dB
  (d) 90 mW
  (e) 11%
9. (c)

CHAPTER 11

1. (a) (0 V, 5 A); (30 V, 0 A)
  (b) 1.5 K ohms
  (c) 16.6 V, 2.3 A, 0.4 V
  (d) 3.8 A p-p; 190

(e) 22.4 V p-p
(f) 0.23 V p-p; 97.5
(g) 18,500; 42.67 dB
(h) 7.66 ohms; No
(i) 100 ohms; No
3. (a) 69 W
    (b) 14.9 W
    (c) 21.6%
5. (a) 2.63%
    (b) 7.4%
7. (a) (0 V, 2.5 A); (40 V, 0 A)
    (b) 1 K ohm
    (c) 20 V, 1.25 A, 1.25 V
    (d) (i) 16.4%
        (ii) 9.2%
9. (i) 22.8% (Current drive.)
    (ii) 7.17% (Voltage drive.)

CHAPTER 12

1. (a) 10.255 mA
    (b) 104.64 mA
3. (a) R_e = 584 ohms; R_B = 148 K ohms
    (b) R_e = 766 ohms; R_B = 118 K ohms
5. (a) 4,325 ohms
    (b) R_1 = 25.12 K ohms, R_2 = 5.2 K ohms
    (c) 4.32 K ohms
    (d) 80 μF
7. (b) 2,800 ohms
    (c) 175
9. (a) 1.82%; 9.54 W
    (b) 1.42% (Actually 0.12% due to negative feedback); 9.32 W

CHAPTER 13

1. (a) 59.5 W
    (b) 39.2 W
    (c) 10.5 W
3. (a) 45 mW
    (b) 11 mW
    (c) 0
5. 0.5° C/W
7. (a) 1.75 W
    (b) 21.4 W
    (c) 25 W
    (d) 60° C
9. (a) 253° C/W
    (b) 0.43 W
    (c) 1.08 W
    (d) 95° C
11 (a) 27.8 W
    (b) 25.5 W
13. (a) Circuit is not inherently stable.
    (b) 376° C/W
    (c) 30.4 W
    (d) 4.93° C/W
    (e) 1.47° C/W
    (f) 69° C
    (g) 43.4 W
15. (a) 25° C
    (b) 900,000 cycles
    (c) 72.5° C
    (d) 3.17° C/W
    (e) 1.7° C/W

CHAPTER 14

(b) 200 K ohms
(d) 175
(e) 2.75 V p-p; 110
(f) R_g=9.52 K ohms; V_g=0.952 V_s; R_L=375 ohms
(g) 175
(h) 1.25×10^{-4}
(j) A_i=-167; R_j=592 ohms; A_v=-106; G_o=1.23×10^{-4} mhos

(k) 0.047; yes
(l) A_i=-175; R_i=600 ohms; A_v=-109
(m) A_{vg}=-6.16; A_{ig}=-165

3. A_i=-160; R_i=40 K ohms; A_v=-224; A_p = 35,800; R_o=25 M ohms; R_o'=56 K ohms

5. (a) 1,070 ohms
   (b) A_i=136; A_v=0.973; R_c=213 ohms
   (c) $h_{oe}R_e=0.0535<0.1$

7. A_i=0.857; R_i=78.6 ohms; A_v=1,635; A_p=1400; R_o=145 K ohms; R_o'=73.7 K ohms; A_{vg}=1450 A_{ig}=0.0966

11. A_i=106; R_i=130 K ohms; A_v=0.979

CHAPTER 15

1. (a) -137
   (b) -188
   (c) 178; Means equations for f_1 in Table 15-1 are valid
   (d) f_o=0.677 Hz; f_p=120 Hz; $f_1 \approx f_p$ = 120 Hz
   (e) 93 Hz
   (f) 127 Hz

3. (a) $f_p=1.4f_o$
   (b) $\frac{A_{v\,(LF)}}{A_{v\,(MF)}} = 0.714 \times \frac{1 + j(f/f_o)}{1 + j(f/1.4f_o)}$
   (c) (i) $0.87\,\underline{\angle 10^\circ}$
       (ii) $0.82\,\underline{\angle 10^\circ}$
       (iii) 0.714

5. (a) 720 Hz
   (b) 1.43 μF
   (c) approximately 100 Hz

7. (a) 12.5%
   (b) 2 Hz
   (c) 90%
   (d) 0.8 V
   (e) 35 μF

9. T1: 1K:3.23K; T2: 24.8K:3.23K; T3: 24.8K:4 ohms

CHAPTER 16

1. (a) 130 mmhos
   (b) 38.5 mmhos @ I_C=1 mA;
       154 mmhos @ I_C=4 mA;
       231 mmhos @ I_C=6 mA
   (c) An increase in temperature would increase I_C, but would also increase V_T, so g_m would be fairly constant.

3. (a) 96 MHz
   (b) (i) $-159\,\underline{\angle -45^\circ}$
       (ii) $-1.01\,\underline{\angle -90^\circ}$
       (iii) $-0.84\,\underline{\angle -90^\circ}$

5. (a) 50 MHz
   (b) 250
   (c) (i) $-5\,\underline{\angle -90^\circ}$
       (ii) $-0.5\,\underline{\angle -90^\circ}$

7. (a) (i) -22
       (ii) 680 kHz
   (b) (i) -44
       (ii) 366 kHz

9. (a) third stage, 550 kHz; second stage, 1025 kHz; first stage, 1270 kHz; overall: 550 kHz
   (b) third stage, 132 kHz; second stage, 547 kHz; first stage, 945 kHz; overall: 132 kHz
   (c) 342 ohms

11. (a) 1.75 MHz
    (b) 290 ohms
    (c) -6.36
    (d) 0.327 μsec

CHAPTER 17

1. (a) 9.1
   (b) 1.5
   (c) -4.5%

3. (a) 16.7; 1.2 MHz

(b) 2×10^7 Hz
(c) –4%
5. (a) 60
(b) 15
(c) 1.34%
7. (a) 8.89
(b) 19.8 K ohms
(c) 4.4 K ohms; $R_o > 40$ K ohms
9. (a) 100
(b) –99
(c) 1980; 19.8
11. (a) $A_v = -23.5$; $R_i \approx 2$ K ohms; $R_i' = 1.34$ K ohms
(b) $k \approx 0.213$
(c) 12.1 K ohms; 12 K ohms; $k > 0.213$ (0.215)
(d) 3.04 K ohms
(e) –3.88; -5; -15.66 dB
13. (a) 76
(b) 66.9 K ohms
(c) 15.4 K ohms
(d) 0.978
(e) 32.9 ohms
(f) 31.7 ohms
15. (a) 1.02 K ohms
(b) –37
(c) 3,200 ohms
(d) 760 ohms
17. (a) 1160
(b) 2.32 M ohms
(c) 266 K ohms
(d) 0.97
(e) 11 ohms

CHAPTER 18

1. (a) +0.33%
(b) +1%
(c) Reduces stability, bandwidth and usually input resistance; Increases distortion and perhaps noise.
(d) +2%
3. (a) 40
(b) Non linear oscillations
(c) –1.25%
5. (a) 6.5 kHz
(b) 65 pF to 0.0065 μF
9. $f_{max,(max)} = 115$ kHz; $f_T = 80$ MHz, $r_{bb'} = 20$ ohms, $C_{b'c} =$ 12 pF
$f_{max,(min)} = 50$ kHz; $f_T = 50$ MHz, $r_{bb'} = 40$ ohm's, $C_{b'c} =$ 20 pF
11. (a) 12 M ohms
(b) 1200 M ohms
(c) 3.08 ohms
(d) 2.14 ohms
13. (a) $\ln 2 \approx 0.7$
(b) 0.42

CHAPTER 19

1. (a) 1.25 W
(b) 5 W
(c) 2½ W; 2½ W
(d) 2½ W; 1¼ W
3. (a) 1.25 W
(b) 2.5 W
(c) 0; 2½ W
(d) 0; 1¼ W
5. (a) 173 ohms
(b) (12 V, 69.5 mA)
(c) $R_1 = 2260$ ohms, $R_2 =$ 544 ohms
7. Yes, $0.17 \times 10^{-5} < 571 \times 10^{-5}$
9. (a) 9%
(b) 3.4%

CHAPTER 20

1. (a) 103 K ohms

(b) A_{v1}=-0.97; A_{v2}=0.98;
V_{s1}=0.97 V p-p, V_{s2}=0.98 V p-p
(c) R_{01}'=1 K ohm, R_{02}'= 26 ohms

3. (a) 17 W
   (b) 19.6 W
   (c) 2.7 W
   (d) 26.6 W
5. (a) 600 ohms: 40 ohms C.T.
   (b) 78%; No, because characteristics do not occupy whole area between I_C and V_{CE} ($V_{CE\,(min)} \neq 0$)
7. (a) 0.257·V
   (b) 3.5 A
   (c) 7 A, 28 V, 24.5 W
   (d) 49 W
   (e) 50%

CHAPTER 21

1. (a) 32 K ohms
   (b) 3000 μmhos
   (c) 96
   (d) N-channel
3. $I_{D,min}$=1 mA, g_{fs}=1300 μmhos
   $I_{D,max}$=16.5 mA, g_{fs}=4000 μmhos
5. (a) -5.82, -8.1
   (b) 2.5 K ohms
   (c) 1,400 M ohms
   (d) 1 M ohm
   (e) 25.2 pF
7. (a) $I_{D,min}$=3 mA, $I_{D,max}$=7.6 mA
   (b) R_S=813 ohms; R_1/R_2=5.8
   (c) If R_2=220 K ohms, R_1=1.28 M ohms. If R_2=1 M ohm, R_1=5.8 M ohms, etc.
   (d) R_S = 1.88 K ohms; R_1/R_2 = 1.42
9. (a) 0.85
   (b) 340 ohms
   (c) 2.2 M ohms
   (d) 0.89: 255 ohms: 11.6 M ohms
   (e) 2.77 pF
11. A_v=27.4; R_i=366 ohms; R_o=280 K ohms

CHAPTER 22

1. (a) P channel enhancement MOSFET
   (b) Device open
   (c) V_{GS} positive
      (i) Good — light goes off when switch closed
      (ii) Open — light off when switch open or closed
      (iii) Shorted — light on when switch open or closed
   (d) V_{GS} positive
      (i) Good — light goes on when switch closed
      (ii) Open — light off when switch open or closed
      (iii) Shorted — light on when switch open or closed
3. (a) V_{DD}=27 V; R_1=2 M ohms, R_2= 106 K ohms
   (b) R_3=11 K ohms, R_4=89 K ohms
   (c) -28
   (d) -7.7
   (e) 11.2 dB
   (f) $A_v \approx$ -3.52; dB decrease = 18
7. (a) -4.9; 5.9 V p-p
   (b) 7.35 V p-p
   (c) -4.1; 6.1 V p-p
   (d) No
9. (a) 2 K ohms
   (b) (i) 39.2 mV
      (ii) 0.59 V
   (c) 378 mW
   (d) (i) 0.53 mV
      (ii) 8 mV
   (e) 108 mW

- (f) 118 V; dissipation capability of the Zener diode
- (g) 23 V

CHAPTER 23

1. (a) −10
   (b) 100 K ohms
   (c) 10,000; 80 dB
   (d) 1 ohm
3. 9.87 V
5. X=2-2/7 V, Y=-13/14 V
7. (a) −96; 100 K ohms
   (b) −98; 100 K ohms
9. (a) 11
   (b) 128 M ohms
   (c) 2.75 ohms
   (d) 101; 45.8 M ohms; 25.2 ohms
11. (a) 10,000
    (b) 2
    (c) 0.40034 V
    (d) 4.40034 V
13. (a) 10.18 V
    (b) -9.91 V
    (c) 10.018 V; -9.991 V
    (d) 20 K ohms
15. (a) (i) −2.27
    (ii) −500
    (iii) 220
    (b) (i) −0.106
    (ii) −500
    (iii) 4,700
    (c) (i) −0.005
    (ii) −500
    (iii) 100,000

CHAPTER 24

3. V_{DD}=-5 V; V_i=0 V, V_o=-5 V;
   V_i=-5 V, V_o=0 V.

CHAPTER 25

1. (a) 5
   (b) 15
   (c) 173
3. (a) 10000
   (b) 10111
   (c) 110101
   (d) 1100101
5. (a) 0010 0011
   (b) 0000 0111
   (c) 0001 0000 0001
   (d) 0101 0111 0010
   (e) 0001 1001 1000 0100
7. (a) Pos.
   (b) Neg.
   (c) Pos.
   (d) Neg.
   (e) Neg.
   (f) Pos.
9. (a) Output = $AB\bar{C} + A\bar{A}\bar{C} + BC\bar{\bar{C}}$
   (b) Output = $AB\bar{C}$
11. (a) $X = \bar{A}\,(\overline{B + C}) \oplus (\overline{CD} + C)$
    (b) $X = \bar{A}\bar{B}\bar{C} \oplus 1$; A = B = C = 0,
    D = 0 or 1
13. $X = \bar{A} + \bar{B}$
17. $X = (A + B)\,(\bar{A} + \bar{B})$
    $= A\bar{A} + A\bar{B} + \bar{A}B + B\bar{\bar{B}}$
    $= A\bar{\bar{B}} + B\bar{\bar{A}}$

CHAPTER 26

3. 4.5 μsec
5. Q = 0, $\bar{Q}$ = 1
7. Q_1 = 0, $\bar{Q}_1$ = 1
   Q_2 = 1, $\bar{Q}_2$ = 0
9. (a) DTL:J=K=1
   (b) RTL:J=K=0
   (c) TTL:J=K=1

11. R_1=36 K ohms, C_1=800 μF
    R_2=12 K ohms, C_2=12 μF
13. (a) 1.65 V
    (b) 2.20 V
    (c) 3.6 V
    (d) 1.73 V
15. (a) Binary count through 7, followed by 000.
    (b) $Q_A, \overline{Q}_B, Q_C; Q_A, Q_B, \overline{Q}_C; Q_A, Q_B, Q_C$
17. (b) (i) 0000 0101 0111
    (ii) 0010 0000 1000
    (iii) 1001 1001 1001
19. (a) 50,000
    (b) 0.5370 V
    (c) Four

CHAPTER 27

1. (a) 3.95×10^{14} to 7.5×10^{14} Hz
   (b) 2.62×10^{19} to 4.97×10^{-19} Joules
   (c) 3.98×10^{-13} Joules
3. (a) 40, 10, 4.4, 2.5 and 1.6 footcandles
   (b) Rectangular hyperbola
5. 44.1, 11, 4.9, 2.76, 1.77 μW/cm^2
7. (a) 0.12 mA
   (b) 2.4 mA
   (c) Between 50 and 100 footcandles
   (d) 1.44 mW, 28.8 mW, 350 mW
9. (0)=1.5 V; (1)=12 V
11. (a) 0.667 mW, 1.23 mW, 0.386 mW
    (b) Optimum R_L = 200 ohms· $P_{L,max}$= 1.6 mW

CHAPTER 28

1. (a) Yes
   (b) V_{GT}=3 V, I_{GT}=0.03A, P_G=0.09 W
   V_{GT}=1 V, I_{GT}=0.17A, P_G=0.17 W
   (c) −40°C No
   +100°C Yes
3. (a) 23.5 ohms
   (b) 100 ohms
5. (a) 6.5 A; 10 μsec; 39 W
   (b) 0.2 A
7. 125 A/μsec
9. (a) 4 A
   (b) 6.4 A
   (c) 2 A
   (d) 0
   (e) 6.3 A
11. (b) (i) $V_C=V_i\ \underline{\angle 7.1^\circ}$
    (ii) $V_C=V_i\times 0.104\ \underline{\angle 84^\circ}$
    (c) R_2=23.2 K ohms; $V_C=0.707\ \underline{\angle -45^\circ}\ V_i$
    (d) (i) 25°
    (ii) Never fires
    (iii) 70°
13. (a) T=0.25 msec; f=4 kHz
    (b) T=0.242 msec; f=4.1 kHz
    (c) R=375 K ohms
    (d) R=37.5 K ohms
    (e) Higher

APPENDIX A

1. (a) 18.5
   (b) 2400 μmhos
   (c) 7.7 K ohms
5. (a) 9 K ohms
   (b) −1 V
7. (a) −2 V
   (b) −18.75; -16.25; -15.75; -13
9. (a) −55.5
   (b) -50
   (c) -85.5
11. 240 K ohms

APPENDIX B

1. (a) 0.34 μF
   (b) 3.4 μF
   (c) 34 μF
3. (b) V_P=185 V, I_P=7.5 mA, V_{GG}=-6 V
   (c) R_e=8.57 K ohms, slope = 1 mA/8.57 V
   (d) (i) –11.25
       (ii) –8.75
       (iii) 2 μF
       (iv) 4.9%
       (v) 3.2%
       (vi) 1.39 W; 1.32 W
       (vii) 35.9 dB
5. (b) 3 K ohms
   (c) –58
   (d) 23.6 dB
   (e) 19.1%
7. (a) 0.973
   (b) 270 ohms
   (c) 0.765
   (d) 24.4 dB

APPENDIX C

1. (a) 3000
   (b) 15; 10
   (c) 1 V p-p
   (d) 69.5 dB; A_1=26 dB, A_2=23.5 dB, A_3=20 dB
3. (a) 2.8 pF; 57 K ohms
   (b) $43.8\underline{/142^\circ}$
   (c) 73.4 pF; 2.17 K ohms
   (d) -j 2.17 K ohms
   (e) V_i=0.002 $\underline{/88.75^\circ}$ volts; V_o=0.088 $\underline{/53^\circ}$ volts
5. (a) –16.1
   (b) –17.2
   (c) 277
7. (a) 4.06 MHz
   (b) 319 kHz
   (c) $9.76\underline{/127.4^\circ}$
   (d) 319 kHz; 27 Hz to 319 kHz
   (e) 1.8 MHz
9. (a) 5.14×10^6 Hz
   (b) 69.4×10^6 Hz
   (c) Same as (a) and (b)
11. 53.4 pF; -22 K ohms; Oscillations possible

APPENDIX D

1. (a) –15
   (b) –13.5
3. (a) (i) 67.8 ohms
       (ii) 67.6 ohms
   (b) (i) 145 V p-p; -73
       (ii) 133 V p-p; -67
       Increased second harmonic distortion
5. (a) 227 ohms
   (b) 60 K ohms
   (c) 26.5 V; -75
   (d) 12.7 V; -36
   (e) 1.09 W; 1.05 W; Grid No. 2 dissipation = 0.38 W
   (f) –78
7. (a) -31
   (b) 74 pF
   (c) 85 pF
   (d) 225 kHz
   (e) $21.9\underline{/135^\circ}$
   (f) 6.97×10^6 Hz
   (g) 450 kHz
9. (b) V_P=250 V, I_P=36 mA, V_{G1}=-15 V, $I_{G2}\approx$ 3 mA
   (c) R_L'=6400 ohms; slope = 20 mA/128 V
   (d) 203 V p-p; 31 mA p-p
   (e) 2%; 0.59 W
   (f) 43.5 dB

INDEX

Accelerometer, optoelectronic, 512
Acceptor impurities, 67-68, 429
Alloy junction, 68
Alloy-type transistor, 155
Alpha, 149-50
Alphanumeric displays, 516-17
Alternator, 27
Aluminum metallization, 431
Amplification:
  concept of, 102
  current, 103
  frequency response, 124-26, 133
  power, 103-104
  voltage, 103
Amplification factor:
  FET, 362
  tube, 548-49
Amplifiers:
  with automatic gain control, 385-87, 390
  bandwidth, 125, 269-77, 284-85, 372, 588
  cascaded transistor stages, 253-56, 258-61, 273-74, 370-72
    bandwidth, 274, 284, 371-72
    high-frequency response, 273-74, 371
    low-frequency analysis, 255-56, 260-61, 371
    mid-frequency gains, 253-55, 258-60, 370
  cascode, 608-609
  cathode follower, 575-76
  chopper, dc, 390-91
  classification (A, AB, B, C), 107-108
  complementary symmetry, 347-50
  differential (see Differential amplifier; Operational amplifier)
  direct-coupled, 582-83
  distortion (see Distortion)
  dynamic range, 135-36
  dynamic transfer curve, 111-13
  emitter follower (see Emitter follower)
  feedback (see Feedback amplifiers)
  frequency response, 124-26, 133, 248-62, 266-77, 370-72, 583-93
  general characteristics, 102-16, 120-36
  graphical analysis of CB configuration, 161, 165, 169
  graphical analysis of CE configuration, 175-85
  grounded-cathode, 552-60, 564-75
  grounded-grid, 576-77
  high 3-dB frequency, 270-77, 371, 588, 604
  input-resistance of, 104-105, 163, 165-66, 177, 237-38, 242-44, 286-87, 289-90, 365, 369-70
  input sensitivity, 135-36
  loudspeaker requirements, 136
  low 3-dB frequency, 248-53, 255-57, 260-61, 371, 587
  maximum power transfer, 106, 259
  noise, 135, 277-79, 609
  notations, v-vii, 103
  operational (see Operational amplifier)
  oscillations, 310-12
  output resistance of, 105, 106, 166, 177, 238, 242-44, 287-88, 291, 365, 368, 370
  pentode, 602-608
  phase relationships, 128, 169-70, 180-81, 334, 340
  phase response, 128
  polar plot, 128
  power (see Power amplifiers)
  power rating, 135
  push-pull, 333-36, 339-50
  radio-frequency, 604-606
  RC-coupled, 248-58, 273-75, 586-88
    high-frequency response, 273-74, 371, 587
    low-frequency response, 248-58, 371, 587
  rejection, 114, 115
  rise time, 274
  square-wave testing, 129
  step response, 256-58, 261-62, 274, 277
  summing, 402-403
  tetrode, 599
  tone controls, 132-34
  transformer-coupled, 211, 220, 258-62, 275-77, 327, 336, 339-47, 604-608
  transistor stage, 142-43, 147-48, 151-55
    common-base, 161-70, 242-44
    common-collector, 155, 242-44, 292-93
    common-emitter, 175-88, 242-44, 248-62, 266-76
      comparison of, 186, 188, 241-44
      current gain, resistive load, 163-64, 166-67, 177, 237, 253-55, 258-60, 270-72
      effect of emitter bypass capacitor, 248-52, 256
      gain-bandwidth product, 269-70, 372
      h-parameter analysis, 237-41
      low frequency response, 248-62, 371
      power gain, 164, 168, 177, 238, 255, 260
      short-circuit current gain, 268-70
      short-circuit current-gain bandwidth product, 269-70
      source impedance
        effect on distortion, 170, 180-85
        effect on gains, 238-42
        effect on response, 248-53
      voltage gain, 164, 167-68, 177, 248-52, 253-55, 260
  triode, 552-60
  tuned, 604-606
  upper 3-dB frequency, 268-77, 371, 588
Amplitude distortion, 108-15, 170, 180-85, 331-32
Analog computation, 403-404
AND gate, 448-49
Angstrom unit, 500
Atom
  Bohr model, 63
  energy-level diagrams, 64
  ionization potentials, 65
Attenuator
  T-network, 127
  voltage-controlled, 389
Automatic gain control, 385-87, 390
Avalanche breakdown, 74
Avalanche diode, 81-87
  dynamic resistance, 82
  temperature characteristics, 85-86
  temperature compensated, 86-87
Avalanche multiplication, 74, 85
Avalanche photodiode, 507

Band-reject filter, 115
Bandwidth, 125, 269-77, 284-85, 372
Barkhausen criterion, 305
Barrier capacitance, 74-76
  potential, 70
Base diffusion, 429
Base-spreading resistance, 267
Base width, 152-53
Bass tone control, 132-34
Beta, 145-46, 149-50
Bias:
  collector-to-base, 195-96
  compensation, 203-204, 348-49
  FET, 366-77
    against device variation, 367-68

Bias (cont'd)
  fixed, 175 76 194
  self-, 196 202, 564 566
  stability, 192 204
  transistor high input resistance circuits, 294 98
  voltage divider with emitter, 199 202, 330, 345 46
Bias line, 367, 566 67
Binary-coded decimal (BCD) system, 467
Binary counters (see Counters)
Binary number system 446
Bistable multivibrators (see FLIP-FLOPS)
Black box theory 230
Bleeder resistor, 44
Bohr atom, 63
Boltzmann constant, 609
Bond, covalent, 65
Bonding wire, 431
Boolean algebra, 448 50, 462 64
Bootstrapping, 297 98, 348, 369
Breakdown diode, 81 87
Breakdown voltages
  in diodes, 7, 74, 81, 85
  in transistors, 147
Breakover voltage, 525-26
Brushless dc motor, 510 12
Bridge amplifier, 414-15
Buried layer, 429

Candle power, 502
Capacitance–
  barrier, 74-76
  diffusion, 75
  space charge 75
  transition, 75
Capacitor
  bypass electrolytic, 202 248 52, 565
  coupling, 252 53, 255 56, 570-71
  series resistance, 41
Capacitor filters, rectifier, 35-41
Carriers in semiconductors, 67 71
Cascaded amplifiers (see Amplifiers, cascaded transistor stages)
Cascade voltage doubler, 54-57
Cascode amplifier, 608-609
Cathode follower, 575-76
Cathode ray tube displays, 76, 111 13, 185-87
Character generators, 516-17
Characteristic curves
  1N759 (Zener diode), 87
  1N2615 (silicon diode), 14
  SK3014 (germanium transistor), 207
  2N2338 (silicon transistor), 160
  2N3054 (silicon transistor), 191
  2N3668 (SCR gate), 528
  2N3903 (silicon transistor), 173
  3N153 (MOSFET), 396
  5Y3-GT (vacuum diode), 14
  6AQ5 (beam power), 607
  6AU6 (pentode), 601
  6CG7 (triode), 548
  6CY5 (tetrode) 599
  12AU7 (triode) 579
  12AX7 (triode), 580
  40319 (silicon transistor), 182
  40421 (germanium transistor) 190
  FPT120 (phototransistor), 509
  GB32J2 (thermistor), 93
  MRD500 (photodiode), 507
Child's Law, 6
Choke input filter 41 44
Circuit notation conventions, v vii, 103
Clamping circuits, 50-54
Classes of amplifier operation 107-108
CLC filter, 45
Clock pulses, 473
Color temperature, 500
Colpitts oscillator, 308 309
Common-base configuration, 147 52, 161 70 239-44
Common collector configuration, 155, 242-44 292 93 (see also Emitter follower)
Common emitter configuration, 142 47, 152 54, 175 86, 231-34, 241 44, 248 61, 266 75, 289 94 328 31
Common mode rejection ratio, 409 12
Common mode signal, 407, 409
Comparator
  digital, 482
  operational amplifier, 416
Compensation, bias, 203 204, 348 49
Complementary MOS, 438-40
Complementary symmetry transistors, 347-50
Conduction angle 529 32
Conduction band, 64 69
Conductivity, in semiconductors, 66-69
Contact potential, 70
Conversion
  ac to-dc, 16 28, 529-31
    (see also Rectifier circuits)
  BCD to-seven segment, 517
Conversion efficiency, 106 107, 168-69, 178-80 327, 341 43, 574
Counters
  applications, 486-90
  cascaded, 486 87
  decade, 485 86
  ripple, 483 85
Covalent bond, 65
Critical inductance, 43-44
Crossover distortion, 340
Crystal oscillator, 387 88
Current drive, 180 84
Current gain, 103, 145 46, 148-50, 177, 237, 242-44, 253 55, 258-60, 268-72
Current gain-bandwidth product, 269-70
Current-regulating diode, 391 93
Curve tracer, 187
Cutin, capacitor filter, 38-40
Cutin voltage
  diode, 7, 9, 18, 73
  transistor, 144 192
Cutoff frequency 268 69
Cutout capacitor filter, 39, 41

D type FLIP FLOP 477-78
Darlington circuit, 294 98
  bootstrapped, 297-98
Dc current gain of transistor, 145-46, 148-50
Dc motor, brushless, 510-12
Decade counters 485-87
Decibels, 120 24
  computation, 120 22
  definition, 120
  frequency response curves, 124 26
  sound levels in 131 32
  voltmeter measurement of 123-24
    impedance correction factor, 124
    range correction factor, 123
  zero reference levels, 122
Decoders
  BCD to-decimal, 487
  lamp driver, 487, 517
  seven segment, 517
Decoupling filters, 310 12
Delay angle 529 30
Delay distortion, 116
Delay time, 314
DeMorgan's Law, 451 52
Depletion region, 69-71, 74, 75
Diac, 534-35
Dicing, 431
Differential amplifier, (see Operational amplifier), 407-19
  applications, 413 19
  common mode rejection ratio, 409 12
  common mode signal, 407, 409
  with constant current source, 412
  gain, 407, 410
  input impedance, 407-413
  in logic circuits, 460
Differentiator, 479
Diffused resistors 430 33
Diffusion, 428-30
  base, 429
  capacitance, 75 76
  emitter, 430
  of impurities, 68, 429
  isolation, 428
Diffusion-type transistor, 156
Diode
  avalanche (see Avalanche diode)
  breakdown, 85
  clamping circuits, 50 54
  construction, 2-3
  dynamic testing of, 76
  field-effect, 391-393
  forward-bias, 4, 70-71

Diode (cont'd)
  four-layer (see PNPN)
  ideal, 4
  integrated, 435
  laser, 516
  light-emitting, 515-17
  peak current requirements, 39-41
  photo-, 504-507
    multiple-junction, 507-509
  PN junction, 68-77
    characteristics curves, 7
    construction, 68
    contact potential, 70
    cutin voltage, 7, 9, 18, 73
    depletion region, 69-71, 74-75
    diffusion capacitance, 75
    dynamic curve, 18-19
    dynamic resistance, 8-9
    electron-hole recombination, 69
    equivalent circuit, 9-11
    forward-biased, 70-71
    maximum temperature of, 73-74
    monolithic, 435-37
    potential profile, 70
    reverse-biased, 71, 74-75
    reverse saturation current, 7, 71-74
    space charge capacitance, 75
    static curve, 16, 18
    static resistance, 8-9
    temperature dependence of, 71-74
    transition capacitance, 75
    Zener voltage, 7
  PNPN, 524-36
    analysis, 526-28
    construction, 525
    gate characteristics, 528-29
    rate effect, 527
    ratings, 530-33
    volt-ampere characteristics, 525-26
  rectifiers, 10-11, 16-27
  reference, 86-87
  reverse-biased, 4, 71
  selenium, 3-4
  Shockley, 534
  surge current, 41
  tunnel, 89-91
  varactor, 75
  Zener, 81-87
Dissipation
  calculation of, 213
  heat sink, 215-17
  hyperbola, 147, 210-11
  maximum rated, 209-11
  plate, quiescent, 210-11
  transistor, 209-24, 326-27, 344-45
Distortion:
  amplitude, 108-15, 170, 182-85, 331-32
  analyzer, 114-15
  compensation, 170, 181, 184
  crossover, 340
  current drive, 180-81, 182-84
  delay, 116
  five point calculation of, 110-11, 182-85
  frequency, 116
  harmonic, 108-15, 170, 182-85, 331-32
  intermodulation, 115
  measurement of, 114-15
  overcompensation, 185
  second harmonic, 108-110
  three point calculation of, 108-110, 182-85
Donor impurities, 66-67
Doubler, voltage, 54-58
Dual matching, 240-41
Dynamic characteristics, 76
Dynamic curve, 17-19
Dynamic impedance, 82
Dynamic load line, 164-65, 178
Dynamic range, 135
Dynamic testing
  diode, 76
  transistor, 185-87
Dynamic transfer curve, 111-13

Early effect, 153
Efficiency, conversion, 106-107, 168-69, 178-80, 327, 341-43, 574
  loudspeaker, 126
EIA, 122, 135
Electromagnetic spectrum, 498-99
Electron emission
  photoelectric (see Photoelectric effects in semiconductors)
  thermionic, 2
Electron hole pair, 65-68
Electronic calculator, 440-42
Electronic charge, 63
Electronic mass, 63
Electro optical shaft encoder, 510-11
Electron shell, 63
Electron volt, 65
Emitter diffusion, 430
Emitter follower, 292-93
  Darlington, 294-98 (see also Common collector configuration)
Energy bands in solids, 64
Energy levels, 64, 66
Epitaxial growth, 427
Epitaxial-type transistor, 157
Equalization, RIAA, 134
Equivalent circuit
  crystal, 387-88
  current regulating diode, 392
  diode, 9-10
  FET, 364
  Norton's, 230, 392
  power supply, 21
  SCR, 526
  Thevenin's, 21, 200, 230
  transformer, 259, 261, 276
  transistor (see Transistor model)
  triode amplifier, 559
  unijunction transistor, 536-38
  varactor, 75
Esaki diodes, 89-91
EXCLUSIVE-OR circuit, 450-52

Failure rate, 223
Fan out, 449
Feedback amplifiers, 282-98
  advantages of, 284-85
  analysis method, 282-83
  bandwidth improvement of, 284-85
  current series, 289-92
  current shunt, 288
  distortion, 285-86
  frequency response, 284-85
  input resistance, 286-87, 289-90
  loop gain, 288, 302
  lower 3 dB frequency, 284-85
  negative, 283-98
  noise, 286
  operational (see Operational amplifier)
  output resistance, 287-88, 291
  positive, 283-84, 304-305
  regenerative, 319, 527
  stability, 284
  upper 3 dB frequency, 284-85
  voltage series, 292-93
  voltage shunt, 293-94
Fermi energy level, 65, 69
Fiber optics, 518
Field effect diode, 391-93
FET (see Field effect transistor)
Field effect transistor
  agc amplifier, 385-87
  applications of FETs, 387-91
  biasing, 366-69, 385-87
  channel ohmic region, 360-61, 387-89
  common drain amplifier, 368-70
  common gate amplifier, 370
  common source amplifier, 364-68
  construction, 357, 379-80
  depletion MOSFET, 378-79
  drain resistance, 361
  dual gate MOSFET, 382-83, 385-87
  enhancement MOSFET, 380-81
  frequency response, 370-72
  gate protection, 383-84
  handling precautions, 381
  insulated-gate, 378-87
  junction type, 356-72
  ON resistance, 388
  operation of, 359-60
  parameters, 361-63
  pinch off voltage, 359-61
  small-signal model, 364-66
  static characteristics, 358, 379-80
  testing, 384-85
  threshold voltage, 380-81
  transfer characteristics, 363, 379-80
  as a voltage variable resistor, 387-91
Filters,
  band-reject, 115

Filters (cont'd)
  decoupling, 310-12
  rectifier, 33-46
    capacitor, 35-41
    inductor, 33-35
    LC, 41-44
    Multiple LC, 44-45
    π-type, 45
    summary table of, 46
  Wien bridge, 115
Fletcher and Munson, 129
FLIP-FLOPS
  clocked, 473-78
  D-type, 477-78
  J-K, 474-77
    master-slave, 477
  race-around, 477
  set-reset, 471-74
  T-type 478
Foot-candle, 500-502
Fourier series, 20
Four-layer diode (see Diode, PNPN)
Frequency
  alpha cutoff, 269
  beta cutoff, 268-69
  lower 3-dB (low) 248-53, 255-57, 260-61, 371, 587
  upper 3-dB (high) 268-77, 371, 588
Frequency distortion, 116
Frequency division, 485
Frequency nomogram, 540
Frequency response, 53-54, 124-26, 133, 248-62, 266-79, 370-72, 586-93, 604

Gain:
  current, 103, 145-46, 148-50, 163-64, 166-67, 177, 237, 242-44, 253-55, 258-60, 268-70
  power, 103, 120-22, 134, 168, 177, 238, 255, 260, 574
  voltage, 103, 164, 167-68, 177, 238, 242-44, 248-55, 260-61, 272-77, 364-65, 368, 370-72, 554-59, 570-73, 602, 609
Gain-bandwidth product, 269-70, 372, 588-91
Gas-filled tubes 87-89
Gates:
  logic, 446-64
    AND, 448-49
    EXCLUSIVE-OR, 450-51
    INVERTER, 438-40
    NAND, 452-56, 457-60
    NOR, 456-61
    NOT, 450-51
    OR, 449, 460-61
Germanium:
  conduction in, 66
  crystal structure, 64-65
  properties, 65-66
  spectral response, 506
Grounded cathode amplifier, 552-60, 564-75
Grounded grid amplifier, 576-77
Ground-noise rejection, 413-14

h-parameters (see Hybrid parameters)
Harmonic distortion, 108-115, 170, 182-85, 331-32
Hartley oscillator, 309-310
Hearing response, 129-32
Heat sinks, 213-20
High-pass filter, 257
Holding current, 525
Holes in semiconductors, 65-71
Hybrid integrated circuits, 437
Hybrid model, 229-44
Hybrid parameters, 229-37
  analysis with, 237-41
  conversion formulas, 237
  obtaining the, 231-33
  typical, 233-36
  variations, 235
Hybrid-π transistor model, 266-68
Hysteresis, 482-83

IHFM rating, 135
Illumination, 500-502
Impedance matching, 259-60
Inductance
  critical, 43-44
  filter, 33-35
  primary, 260-62
  swinging choke, 43
Infrared emitters, 516
Input capacitance, 270-72, 365-66, 370, 584-85
Input characteristics of transistors
  common-base, 148-162
  common-emitter, 144, 176
Input impedance of voltmeter, 54, 57
Input resistance, 104-105, 163, 165, 170, 177, 237-38, 242-44, 286-87, 289-90, 365, 369-70
Input sensitivity, 135-36
Input waveforms, 170, 180-81
Integrated circuits
  components, 432-35
  fabrication, 426-42
  linear, 400-20
  nonlinear, 446-64, 471-90
Integrator, 403-404
Interbase resistance, 538
Intermodulation distortion, 115-16
Intermodulation frequencies, 115
Intrinsic semiconductor, 65
Intrinsic stand-off ratio, 538
Inverse peak-voltage rating, 6, 25, 41, 46, 74, 81
INVERTER circuit, 314, 438-40, 450-51
Ionization, 65
Ionization potential, 65
Isolation diffusion, 436
Isolation islands, 429

J-K flip-flops, 474-77
Junction diode (see Diode, PN junction)
Junction-type photoconductors, 504-509

Large-scale integration, 440-42
Large-signal amplifiers (see Power amplifiers)
LASCR, 525
Laser diode, 516
LC filter, 41-44
LEDs 515-17
Light
  fundamentals of, 498-503
  operated relays, 509-510
  photoemission (see Photoelectric effects)
  photon nature of, 498-500
  sources, 500, 515-17
  velocity, 498
Light dimming circuit, 535
Light-emitting diodes, 515-17
Linear integrated circuits, 400-420
Liquid crystal display, 517
Load line
  ac, 164-65, 178, 567-71
  analysis, 16-18
  dc, 161-63, 175-76, 552-53
  dynamic, 164-65, 178, 567-71
  equation of, 17, 161, 175, 553
  slope of, 17, 164-65, 178, 567-69
  push-pull, 342
  static, 161-63, 175-76, 552-53
Loftin White, 582-83
Logic
  circuit packaging, 431-32
  combinational, 446-64
  complementary MOS (CMOS), 438-40, 459-60
  current-mode (CML), 460-61
  diode (DL), 448-49
  diode-transistor (DTL), 347
  emitter-coupled (ECL), 460-61
  high-threshold (HTL), 454-55
  levels, 447
  MOS, 457-58
  negative, 447
  positive, 447
  pulse, 448
  resistor-transistor (RTL), 456-57
  sequential, 471-90
  transistor-transistor (TTL), 455-5
  truth tables, 448-53, 457, 472-76, 484
  wired, 469
Logic gates (see Gates, logic)
Loop gain, 288, 402
Loudness levels, 129, 131-32
Loudspeaker requirements, 136
Low-pass filter, 274
Lumen, 500, 502

Majority carrier, 67-71
Masking, diffusion, 429
Master-slave FLIP-FLOPS, 477

Matching impedance, 259-60
Maximum power gain, 240-41
Maximum power transfer, 106, 126, 259
Medium-scale integration, 440
Memory:
    one-bit, 478
    read-only, 441-42
Mercury-arc rectifier, 88
Metallization, 431
Metal-oxide-semiconductor (MOS)
    capacitor, 434
    complementary symmetry, COS/MOS, 438-40
    transistor (see Field-effect transistor)
Micron, 498
Microphonics, 609
Miller's theorem, 270-71, 585
Minority carrier, 67-71
Mobility, 68
Monolithic IC, 427-32, 436
Monostable multivibrator, 478-80
MOS (see Metal-oxide-semiconductor)
Motorboating, 311
Multiplexers, 612-13
Multiplier, voltage, 58-59
Multivibrators
    astable, 315-20
    bistable (see FLIP-FLOPS)
    monostable, 478-80
Music power rating, 135
Mutual conductance, 362, 549-51

N-type impurities, 66-67
NAND circuit, 452-56, 457-60
Negative resistance:
    in triode amplifiers, 591-92
    in tunnel diodes, 90-91
    in unijunction transistors, 538
Noise, 277-79, 609
Noise bandwidth, 609
Noise figure, 135, 277
Nomogram, frequency, 540
Nonlinear distortion, 108
    (see also Distortion)
Nonlinear integrated circuits, 446-64, 471-90
NOR circuit, 456-61
NOT circuit, 450-51
Notation, general amplifier, 103
Notation, transistor circuit, (v-vii), 142-43

One-shot, 478-80
Operating point in transistor circuits, 162-63, 175-76, 193, 195, 198, 210, 221, 400-19
Operational amplifier (OP AMP) 400-19
    adder, 402-403
    in analog computation, 403-404
    applications, 413-19
    common-mode rejection ratio, 409-12
    common-mode voltage, 407, 409
    comparator, 416
    differential gain, 407, 410-11
    general purpose 408-409
    high performance, 412-13
    input impedance, 402, 405-406, 413
    integrated, 408, 412
    integrator, 403
    inverting, 401-402
    loop-gain, 402
    noninverting, 406-407
    virtual ground, 402
Optical paper tape reader, 510
Optoelectronic devices, 498-518
Optoelectronic spring-mass accelerometer, 512
OR gate, 449, 460-61
Oscillators
    relaxation, 91, 538-39
    sinusoidal, 304-10
        Barkhausen criterion, 305
        Colpitts, 308-309
        crystal, 387-88
        Hartley, 309-310
        Pierce, 387-88
        phase-shift, 306-307, 387
        resonant circuit, 307-310
        Wien bridge, 307, 416-17
Oscilloscope display.
    amplifier's dynamic transfer curve, 111-13
    diode's dynamic characteristics, 76
    PN junction capacitance, 76
    transistor characteristic curves, 185-87
Output power meter, 126-28
Output resistance, 105-106, 177, 238, 242-44, 365, 368, 370

P-type impurities, 67, 68, 429
Parameters:
    field-effect transistor, 361-63
    transistor (see Transistor)
    triode, 548-51
Partition noise, 609
Peak current requirements, 39-41
Peak inverse voltage, 6, 25, 41, 46, 74, 81, 525
Phanotron, 88
Phase control, 530, 534-35
Phase distortion, 116
Phase-inverter circuits, 334, 341
Phase relationships, 128, 169-70, 180-81
Phase-shift measurement, 613-14
Phase-shift oscillator, 306-307, 387
Phons, 129, 131
Photoconductive cells, 503-504
Photoconductors, 504-508
Photocouplers, 517-18
Photodiode, 504-507
Photoduodiode 507
Photoelectric effects in
    semiconductors, 503, 506
    spectral response, 504, 506, 515
Photoemissive sensors, 515
Photoexcitation, 503, 506
Photo-FET, 508
Photomultiplier tube, 515
Photons, 498-500, 503
Photoresist, 429
Photoresistors, 503-504
Photo-SCR, 508-509
Phototransistors, 508-509
Photovoltaic effect, 513-15
Pierce oscillator, 387-88
Pinch-off voltage, 359-61
Pinch resistor, 433
Planck's constant, 498-500
PN junctions, (see Diode, PN junction)
PNPN diode (see Diode, PNPN)
Potential hill (barrier), 70
Power amplifiers
    class A, 326-36
    class AB operation, 340-41
    class B, 339-50
    complementary-symmetry, 347-50
    conversion efficiency, 327, 341-43
    integrated, 419-21
    push-pull, 333-36, 339-50
    transformer-coupled, 327-33, 606-608
    transformerless, 347-50
Power control circuits, 529-31, 534-35, 540
    conduction angle, 529-31
    delay angle, 529-30
    full-wave, 530-31
    half-wave, 529-30
    phase control, 530, 534-35
Power derating:
    thyristors, 531
    transistors, 209-10
Power gain:
    decibel, 120-24
    definition, 103
Power ratings (see Ratings)
Power supply (see Rectifier circuits; Regulated power supply)
Power-temperature derating curves, 209-210, 531
Pre-ohmic etch, 430
Propagation delay in logic circuits, 455-462
Pull-up:
    active, 456
    passive, 456
Pulse amplifying capability, 261-62
Push-pull amplifier, 333-36, 339-50

Quiescent point (Q point), 193, 195, 198, 210-11, 221

- Race around condition, 477
- Radio frequency interference, 535-36
- Rate effect, 527
- Ratings
  - amplifiers, 135, 136
  - diodes, 74, 82
  - thermal cycling, 222-24
  - thyristors, 530-33
  - transistors, 147-209
  - transformers, 23-24
- Ratio of rectification, 22, 25-35
- Recombination in semiconductor, 69
- Recording characteristic, 134-35
- Rectifier
  - mercury arc, 88
  - semiconductor (see Diode, PN junction, Rectifier circuits)
  - silicon controlled, 524-32
  - vacuum tube, 2-10
- Rectifier circuits, 16-28
  - bridge, 26
  - full wave, 24-26
  - half-wave, 10-11, 19-24
  - regulation in, 20
  - shunt, 50-54
  - three phase, 26-27
  - voltage-doubling, 54-58
  - voltage multipliers, 58-59
- Rectifier filters (see Filters, rectifier)
- Reference Zener diodes, 86
- Regulated power supply, 417-19
  - integrated, 418
  - series voltage regulator, 417
  - shunt voltage regulator, 417
- Relaxation oscillator, 538-39
- Resistance, negative (see Negative resistance)
- Resistivity, in semiconductors, 68
- Resonant circuit oscillator, 307-10
- Reverse breakdown voltage, 374, 525
- Reverse saturation current
  - PN diode, 7, 71-74
  - transistor, 145, 150, 154, 192-93
- RIAA equalization, 134
- Ripple counters, 483-85
- Ripple factor, 21-22, 25-27, 34-39, 42-46
- Rise time, 274, 314

- Sag, 257, 261-62
- Saturation region
  - common-base configuration, 162
  - common emitter, 175, 313-14
- Saturation resistance, 313-14
- Schmitt trigger, 480-83
- Selenium photovoltaic cell, 514-15
- Self-bias, 196-202
- Semiconductor numbers, 12
- Semiconductors
  - acceptor impurities, 67-68
  - diode (see Diode, PN junction)
  - donor impurities, 66-67
  - doping, 66-68
  - energy bands in, 64-66, 69
  - Fermi level in, 65-69
  - intrinsic, 65
  - majority carriers in, 67-71
  - minority carriers in, 67-71
  - N type, 66-67
  - P type, 66-67
  - photoelectric effects in, 503-506
  - PN junction (see Diode, PN junction)
  - recombination in, 69, 152-53
  - transistors (see Transistor)
- Sensitivity
  - amplifier, input, 135-36
  - of photodiode, 506
  - of phototransistor, 508
  - spectral, 504
- Seven segment display, 487, 517
- Sheet resistance, 433
- Shockley's relation, 8
- Short circuit current gain, 268-70
- Shot noise, 609
- Shunt rectifier, 50-54
  - in a VTVM, 52-54
  - with filter, 52
- Signal to noise ratio, 135, 277
- Silicon
  - photovoltaic cell, 514-15
  - spectral response, 514-15
- Silicon controlled rectifier, 524-32, (see also Diode, PNPN)
  - in power control circuits, 529-31
- Silicon controlled switch, 525
- Silicon diode (see Diode, PN junction)
- Solar panels, 514
- Sound level meter, 132
- Sound levels, 131
- Space charge, 5
- Space-charge capacitance, 75
- Space charge region in PN junction, (depletion region), 69-71, 153
- Spectral response, 500-502, 504, 506, 515
- Square wave response, 256-58
- Square-wave testing, 129-30
- S R FLIP-FLOPS, 471-77
- Stability
  - bias, 192-204
  - of common base, 194
  - of feedback amplifier, 284
  - of operational amplifier, 412
  - thermal, 192-204, 209-24
- Stability factors, 193-203
- Stabistor, 87
- Static characteristics (see Characteristic curves)
- Step response, 256-58, 261, 274-75, 314-15
- Storage time, transistor, 314
- Surge current, 41
- Swinging choke for filters, 43

- T type FLIP-FLOPS, 478
- Temperature coefficient
  - expansion, 223
  - thermistors, 91
  - Zener diodes, 85-86
- Thermal fatigue, 222-24
- Thermal noise, 609
- Thermal resistance, 211-20
- Thermal runaway, 193, 210-11, 332-33
- Thermal stability, 192-204, 209-24
  - conditions for, 220-22
- Thermionic emission, 2
- Thermistor, 91-96
- Thevenin's equivalent circuit, 21, 200, 2
- Thin-film capacitors, 434
- Thin-film resistors, 434
- Three-phase rectifier, 26-27
- Threshold voltage, 380-81
- Thyratron, 88
- Thyristors, 524-36, (see also Diode, PNPN)
- Tilt, 257, 261-62
- Tone controls, 132-34
- Totem-pole configuration, 456
- Transconductance.
  - FET, 361-62
  - transistor, 267
- Transfer characteristic.
  - amplifier, 109, 111-13
  - FET, 363, 379-80
  - transistor, 140
  - tube, 551
- Transformer-coupled amplifier, 258-62, 275-77, 327-36, 339-47
- Transformer impedance matching, 259-60
- Transformer utilization factor, 23-24, 26, 35
- Transistor
  - ac current gain, 145-46, 149-50, 163-64, 166-67, 177
  - alloyed junction, 155
  - alpha, 149-50
  - amplifier equations, 237-41
  - base-spreading resistance, 267
  - base width, 152-53
  - basic characteristics, 142-57
  - beta, 145-46, 155
  - biasing circuits, 176, 192-204
  - breakdown voltages, 147
  - characteristics
    - analytic expressions for, 146, 150
    - collector, CB, 148
    - collector, CE, 144
    - cutoff region, 145, 154, 161, 175, 312-13
    - input, CB, 148
    - input, CE, 144
    - saturation region, 162, 175, 313-14
  - collector cutoff current, 145, 150, 154
  - common-base configuration, 147-52, 161-70
  - common-collector configuration, 155
  - common-emitter configuration, 142-47, 152-54, 175-88, 248-62, 266-75, 289-94, 327-34

- Transistor (cont'd)
  - complementary symmetry, 347-50
  - construction, 155-57
  - current components, 143, 152-53
  - current gain, 145-46, 148-50, 163-64, 166-67, 177
  - curve tracer, 187
  - cutoff region, 145, 150, 152, 153-55, 161, 175-76, 312-13
  - dc current gain, 145-46, 149-50
  - delay time, 314
  - diffusion type, 156
  - diode analogy, 143
  - dissipation, 209-25, 326-27, 344-45
  - drift-field type, 156
  - epitaxial type, 157
  - fall time, 314-15
  - field effect (see Field effect transistor)
  - forward current transfer ratio CB, 148-50
  - forward current transfer ratio, CE, 145-46, 232
  - graphical analysis, 161-65, 169, 175-85
  - grown-junction type, 155
  - hybrid model, 229-44
  - hybrid – π model, 266-68
  - integrated, 435-37
  - junction operating temperature, 202, 209
  - junction voltage values, 144, 192
  - lead identification, 157
  - leakage current, 154-56
  - maximum frequency of oscillation, 310
  - mesa type, 156
  - model hybrid, 229-44
    - hybrid – π, 266-68
    - two-diode, 143
  - noise, 277-79
  - parameters.
    - alpha, 149-50
    - beta, 145-46, 155
    - capacitances, 266-67
    - conductances, 267
    - display of, 187
    - hybrid, 231-36
    - hybrid – π, 267-68
  - photo-, 508-509
  - planar, 156
  - ratings, 147, 209
  - reverse saturation current, 145, 150, 154, 192-93
  - rise time, 314-15
  - saturation region, 162, 175, 313-14
  - saturation resistance, 314
  - short-circuit current gain, 268-70
  - storage time, 314
  - temperature variation:
    - of dc current gain, 146, 192
    - of hybrid parameters, 235
    - of hybrid – π parameters, 267
    - of junction voltages, 192
    - 150, 192-93
  - testers, 151
  - transfer characteristic, 140
  - turn-off time, 314-15
  - tur
  - voltage gain, 143, 164, 167-68, 177
- Transistor amplifier (see Amplifiers)
- Transition capacitance, diode, 76
- Transition region, 70-71
- Treble tone control, 132-34
- Triac, 532-33
  - in power control circuits, 534-36
- Triggering devices, 533-34, 540
- Truth tables, 448-53, 457, 472-76, 484, 517
- Tuned amplifiers, 604-606
- Tunnel diode, 89-91
- Tunneling phenomenon, 90

40

- Vacuum tube
  - beam power, 606-608
  - diode, 2-10
  - pentode, 599-608
  - tetrode, 598-99
  - triode, 546-60, 564-77, 582-93
  - type numbers, 11
- Valence band, 63-68
- Varactor diodes, 75
- Voltage amplification, 103, 164, 167-68, 177, 238, 242-44
- Voltage comparator, 416
- Voltage-controlled attenuator, 389-90
- Voltage-doubling rectifier, 54-58
  - full-wave, 57-58
  - half wave, 54-57
  - in a VTVM, 55-57
- Voltage drive, 181-85
- Voltage gain, 103
- Voltage multiplier, 58-59
- Voltage regulation, 20-21, 42
- Voltage regulator
  - cold-cathode, 88
  - design of, 83-85
  - field-effect diode, 393
  - IC op amp., 417-19
  - series, 417
  - shunt, 417
  - thermistor (ac), 95
  - Zener diode, 82-85
- Voltage variable resistor, (VVR), 387-91
- Volt-ampere characteristics (see Characteristic curves)
- Voltmeter
  - ac readings, 27-29, 54
  - FET, 391-92
  - frequency response, 53, 57
  - VTVM Heathkit, 55-57
  - VTVM, Hewlett Packard, 52-54
- Waveform considerations, 170, 180-81
- White noise, 278
- Wien bridge oscillator, 307, 416-17
- Wired logic, 469

- y parameters, 361-62

- Zener breakdown, 86
- Zener diodes, 81-87, 435
- Zener voltage, 7, 81
- Zero crossing detector, 536